Building a Quantitative Linguistic Profile of Bridge Team Communication for a Performance Assessment of Native and Non-Native Speakers of Maritime English

Peter John

Certified Translator and Interpreter

Supervisors:

Benjamin Brooks

Ulf Schriever

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of Tasmania, Australian Maritime College

March 2019

Abstract

This doctoral thesis studies Bridge Team Communication from a quantitative, text corpus perspective. This approach seeks to model idiosyncratic linguistic patterns for native and non-native speakers of English.

Working on board sea-going merchant ships provides a unique professional environment with crews being recruited from the most varied cultural and linguistic backgrounds. Multinational and multilingual crews live and work in a confined space, and they co-operate in a complex socio-technical environment by carrying out a large number of different tasks in which communication is a core skill. The shared technical knowledge required for manoeuvring ships safely around the globe needs to be communicated effectively among all participants in a safe and efficient manner using a language every crew member understands. Conversely, miscommunication may erode the team members' situational awareness and this can be a root cause in tragic maritime incidents.

Maritime education and training (MET) institutions have been addressing this issue by actively training future seafarers in realistic scenarios and simulating real-life situations on board. The seafarers with responsibility for navigation of the ship (bridge teams) are trained in simulated environments with a high degree of ecological validity, thus truly reflecting the interaction on board sea-going vessels. By means of specifically defined training exercises, different nautical situations are simulated and relevant communication skills for the varying requirements taught.

Applied Linguistics research opens up a whole range of possibilities to study naturalistic human interaction and language production in nautical training procedures which combine the authenticity of field studies with the controlled environment of research experiments. This doctoral research sets out to develop a quantitative model for bridge team communication as a sub-genre of the sociolinguistic language variety (or *genre*) of Maritime English. The overall aim is to provide a methodology for assessing verbal language performance of native and non-native speakers (in the sense of Saussure's *parole*) in this particular English for Specific Purposes (ESP) domain.

For studying the discourse community of (future) nautical officers, a synchronic text corpus has been developed on the basis of verbatim transcripts of empirical speech

events recorded in the ecologically valid simulators mentioned above. The specialised spoken corpus consists of some 107,000 word tokens and is analysed by means of corpus linguistics methods from a sociolinguistic, psycholinguistic and corpus-pragmatics perspective. Between-groups analyses investigate a series of dependent linguistic variables including vocabulary growth, word frequencies, lexical and key word densities, and part-of-speech diversity as well as linguistic proxy variables for cognitive load. Additionally, risks of miscommunication are studied pragmatically by following Searle's Speech Act theory.

The analyses lead to typical statistical distributions for linguistic variables of the sampled sociolinguistic groups of native and non-native speakers of English. By comparing the observed variables with reference text corpora outside a maritime setting, benchmark values are provided to gauge maritime idiomaticity. The computed statistical distributions also provide a robust methodology for making inferences on the communicative effectiveness of bridge teams and their members. The research fills a gap of empirical research on bridge team communication and creates a quantitative model for determining idiosyncratic differences between the sub-genre of bridge team communication and other communicative settings.

The findings of this research contribute towards an improved methodology for assessing verbal communicative performance of future nautical officers. Especially when considering that a high number of all incidents at sea are caused by deficient or ineffective communication, a linguistic model of bridge team communication will help to advance maritime communication standards, thus leading to a better understanding in a truly multinational work environment.

Declaration of originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Authority of Access

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Statement of ethical conduct

The research associated with this thesis abides by the international and Australian codes on human and animal experimentation, the guidelines by the Australian Government's Office of the Gene Technology Regulator and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University.

Date: 08/03/2019

Peter John

Statement of co-authorship

The manuscripts contained in this thesis comply with the University of Tasmania's Authorship of Research Policy. The following people and institutions contributed to the publication of work undertaken as part of the thesis:

Peter John,

Australian Maritime College, University of Tasmania

= Candidate/Author 1

Benjamin Brooks,

Australian Maritime College, University of Tasmania

= Author 2

Ulf Schriever,

Australian Maritime College, University of Tasmania

= Author 3

Christoph Wand,

Faculty of Maritime and Logistics Studies,

Jade University of Applied Sciences

= Author 4

Author details and their roles

Paper 1 (located in Section *10.1. Paper I*)

John, P., Brooks, B., Wand, C. & Schriever, U. (2013). Information density in bridge team communication and miscommunication – a quantitative approach to evaluate maritime communication. *WMU Journal of Maritime Affairs*, *12(2)*, 229-244.

The Candidate is the main author and was primarily responsible for the conception, planning and execution of the work. Authors 2 and 3 contributed to the idea as well as its formalisation and refinement. Author 4 contributed to the statistical calculations undertaken. The Candidate's contribution is approximately 80%.

Paper 2 (located in Section *10.2. Paper II*)

John, P., & Brooks, B. (2014). Lingua Franca and its Grammar Footprint: Introducing an Index for Quantifying Grammatical Diversity in Written and Spoken Language. *Journal of Quantitative Linguistics*, *21(1)*, 22-35.

The Candidate is the main author and was primarily responsible for the conception, planning and execution of the work, including the quantitative and qualitative data analysis. Author 2 contributed to its formalisation and refinement. The Candidate's contribution is approximately 80%.

Paper 3 (located in Section *10.3. Paper III*)

John, P., Brooks, B. & Schriever, U. (2017). Profiling maritime communication by non-native speakers: A quantitative comparison between the baseline and standard marine communication phraseology. *English for Specific Purposes, 47*, 1-14.

The Candidate is the main author and was primarily responsible for the conception, planning and execution of the work, including the quantitative data analysis. Authors 2 and 3 contributed to the idea as well as its formalisation and refinement. The Candidate's contribution is approximately 80%.

Paper 4 (located in Section *10.4 Paper IV*)

John, P., Brooks, B. & Schriever, U. (2017). Linguistic measurement of cognitive load in maritime team communication by native and non-native speakers of English. Submitted to *Applied Ergonomics*.

The Candidate is the main author and was primarily responsible for the conception, planning and execution of the work, including the quantitative data analysis. Authors 2 and 3 contributed to the idea as well as its formalisation and refinement. The Candidate's contribution is approximately 80%.

Paper 5 (located in Section *10.5 Paper V*)

John, P., Brooks, B. & Schriever, U. (2019). Speech acts in professional maritime discourse: A pragmatic risk analysis of bridge team communication directives and commissives in full-mission simulation. *Journal of Pragmatics, 140*, 12-21.

The Candidate is the main author and was primarily responsible for the conception, planning and execution of the work, including the quantitative data analysis. Authors 2 and 3 contributed to the idea as well as its formalisation and refinement. The Candidate's contribution is approximately 80%.

Statement of co-authorship

We, the undersigned, agree with the above-stated proportion of work undertaken for each of the published, accepted or submitted peer-reviewed manuscripts contributing to this thesis.

Candidate/Author 1:	Date: 08/03/2019
Peter John	
Author 2:	Date: 08/03/2019
Benjamin Brooks	
Author 3:	Date: 08/03/2019
Ulf Schriever	
Author 4:	Date: 08/03/2019
Christoph Wand	
Director, Australian Maritime College - Seafaring	Date: 28/04/2020
A/Prof. Prashant Bhaskar	

Preface

In the year 2010 I visited the Australian Maritime College for the first time, after having worked in maritime education and training for over 14 years and having attempted unsuccessfully to initiate doctoral research activities in the domain of Maritime English. It was during this first visit when I was given the opportunity to enrol as a PhD student in the linguistic domain I have been working in nearly my entire professional career.

My first thanks shall go to my supervisors, Ben Brooks and the late Ulf Schriever, who encouraged me to take this on and have guided me patiently through the process of defining my research, writing academic papers and finalising this dissertation.

I would also like to express my gratitude to John Lloyd who was instrumental in granting me a scholarship. Many thanks to Stephen Cahoon and Marcus Bowles, Hossein (Behrooz) Enshaei and other members and peer students at the AMC with whom I had very interesting discussions on the methodological aspects of my research, and who have given me a pretty good insight into Australian culture.

In addition to the AMC members, many people have contributed significantly to this research. Special thanks to those colleagues who have supported me in recording full-mission simulator sessions. Without the assistance of Christoph Wand, Hans-Jörg Nafzger, Klaus Damm, Christian Sievers and Thorsten Kramer of Jade University of Applied Sciences as well as Bill Kavanagh, Peter Walter and Vivion Gough of the National Maritime College of Ireland this research undertaking would not have been possible. The same applies to those nautical students in Germany and in Ireland who volunteered to being audio-recorded.

Thanks also to other researchers and IMECeers who are active in maritime English and who have assisted me in testing the methods in adjacent fields of nautical communication. Partners in crime indeed, Alison Noble, Peter Björkroth and Naoyuki Takagi.

Finally, I would like to thank my family for their patience and emotional support during all those hours I spent at my desk.

“One thing is gone, to return no more forever – the romance of the sea.”

Mark Twain, 1893

Table of contents

List of figures

List of tables

List of abbreviations

AL	Applied Linguistics
ANOVA	Analysis of Variance
BTC	Bridge Team Communication
DA	Discourse Analysis
ESP	English for Specific Purposes
GDI	Grammar Diversity Index
IALA	International Association of Lighthouse Authorities
IMEC	International Maritime English Conference
IMLA	International Maritime Lecturers' Association
IMO	International Maritime Organization
ITU	International Telecommunications Union
L1	English as a First Language
L2	English as a Second Language
LF	Lingua Franca
LSP	Languages for Specific Purposes
ME	Maritime English
MET	Maritime Education and Training
MWU	Mann-Whitney U test
PDI	Part-of-Speech Diversity Index
POS	Part-of-Speech
PS	Probability of Superiority
QCA	Quantitative Content Analysis
SA	Situation Awareness

SMCP	Standard Marine Communication Phrases
SMNV	Standard Marine Navigational Vocabulary
SOLAS	Safety of Life at Sea
spdi	Special Part-of-Speech Diversity Index
STCW	Standards on Training, Certification and Watchkeeping
VHF	Very High Frequency
VTS	Vessel Traffic Service

List and status of publications

Paper No. I:	"Information density in bridge team communication and miscommunication – a quantitative approach to evaluate maritime communication"
nature:	conceptual / research paper
published in:	World Maritime University Journal of Maritime Affairs (2013) (full-length, double-blind review)
Paper No. II:	"Lingua Franca and its Grammar Footprint: Introducing an Index for Quantifying Grammatical Diversity in Written and Spoken Language"
nature:	conceptual paper / research paper
published in:	Journal of Quantitative Linguistics (2014) (full-length, double-blind review)
Paper No. III:	"Profiling maritime communication by non-native speakers: A quantitative comparison between the baseline and standard marine communication phraseology"
nature:	research paper
published in:	English for Specific Purposes (2017) (full-length, double-blind review)
Paper No. IV:	"Linguistic measurement of cognitive load in maritime team communication by native and non-native speakers of English"
nature:	research paper
published in:	under review
Paper No. V:	"Speech acts in professional maritime discourse: A pragmatic risk analysis of bridge team communication directives and commissives in full-mission simulation"
nature:	research paper
published in:	Journal of Pragmatics (2019) (full-length, double-blind review)

1. Introduction

Sea-going ships are navigated as a co-operative task carried out by professional teams. These *bridge teams* consist of watch-keeping navigational officers and ratings (and sometimes pilots) who communicate verbally in order to exchange the technical information required for steering their vessel in a safe and efficient manner. The verbal information exchange on a ship's bridge is referred to as *Bridge Team Communication* (BTC).

Due to international regulations and the fact that crews on board merchant ships are nearly always multinational, bridge teams are obliged to communicate in English. The sociolinguistic English variety used in the maritime domain is commonly known as *Maritime English* and falls within the category of English for Specific Purposes (ESP).

Maritime English is a compulsory element in Maritime Education and Training (MET), and most shipping companies have effectively made a sufficient Maritime English *proficiency* a mandatory element for employing seafarers[1]. Language proficiency is usually assessed in a typical test environment, be that in a classroom setting or in a computer laboratory.

On the other hand, the notion of *communicative performance* reaches beyond a grammatically and lexically correct language and towards sociolinguistic, discourse and strategic competencies (Canale 1983; Canale and Swain 1980; Hyme 1972; Paltridge 2006). For working in a team on board ships, communicative competence is of utmost importance to avoid damages to ship and cargo, injuries and death to crew members and harm to the environment (Bocanegra-Valle 2011; Cole and Trenkner 2009; de la Campa Portela 2005; Jurkovič 2015; Pritchard & Kalogjera 2000).

The discrepancy between language proficiency and communicative performance is a reality the researching student has regularly noticed in more than 20 years of teaching Maritime English. He has experienced striking differences between students' achievement in *proficiency* assessment and their language *performance* both in full-mission simulation and in real-life situations on board training ships. All too often,

1 See Noble (2017, p. 125): "The latest amendments to STCW 78, as amended (Table A-II/1, Manila Amendments 2010) stipulate that officers "shall" use IMO SMCP and "use English in written and oral form".

Maritime English learners who perform well in written and spoken language proficiency assessment fail to communicate effectively in professional team tasks due to an insufficient communicative performance.

The observed discrepancy between language proficiency and communicative performance has been the motivation for developing a methodology which would provide a valid assessment methodology for naturalistic verbal interaction by bridge team members while they are engaged in navigational tasks. During the performance of these navigational tasks various challenges need to be attended to simultaneously, a fact that often leads team members to perform far worse linguistically than in a language proficiency assessment under optimal assessment conditions without any distraction.

The aim of this doctoral research is to develop a quantitative methodology for a language performance assessment in the maritime domain. With computerised speech recognition on the brink of becoming a standard technology, naturalistic verbal interaction can soon be transcribed and assessed in real-time. This research aims to contribute some of the quantitative methods required for developing computer tools which will provide a valid and reliable means of assessing language performance in a simulated or real professional environment.

This chapter provides background information on the research area, it outlines the legal framework for maritime communication and it categorises the different communicative settings or *speech events* on board sea-going ships. It continues by reviewing existing research and subsequently identifying the research gaps studied in this dissertation. The chapter concludes with the formulation of principal and subordinate research questions which are analysed in the course of this thesis and by the supporting publications in academic journals.

1.1. Research background

Every now and then a major shipping accident hits the headlines in the public media, soon to be followed by an expert analysis of its causes. In many cases, human factors including *miscommunication* are identified as a contributory factor to groundings, collisions, fire on board and other spectacular cases. However, the incidents reported in the mass media are merely the tip of the iceberg: for the period from the year 2005 to 2014, the Casualty Statistics issued by Lloyd's List Intelligence report a total of 24,545 shipping incidents (Lloyd's List Intelligence 2016). The European database on marine accidents includes "a total of 9,180 occurrences" (European Maritime Safety Agency 2015, p. 8) for the period from the year 2011 to 2014, out of which 67 percent are related to "human erroneous actions" (ibid, p. 8). These 6,151 cases mean that on a global scale, human error contributes to an average of four reported maritime incidents per day.

While reported incidents provide unambiguous figures on the frequency of maritime *occurrences*, the situation is less clear for communicative problems involved in these incidents. Estimates suggest that up to 40% of all maritime casualties include some form of *miscommunication* either as their main cause or as a contributory factor[2].

Today, the majority of ocean-going ships are crewed with multinational teams who communicate in English as a second language (L2) in a *Lingua Franca* (LF) communicative context (Noble 2017; Noble et al 2011a, 2011b; Sampson 2013; Schriever 2008, 2005). In the particular Lingua Franca context on board merchant ships, English is used "for communication between speakers who have no native language in common" (Davies 2005, p. 85).

Expressed in figures, in the year 2015 the world merchant fleet consisted of 68,723 ships which were crewed with a supply of 1,647,500 seafarers (774,000 officers and 873,000 ratings). Most crew members originated from China, The Philippines, Indonesia, the Russian Federation and Ukraine (International Chamber of Shipping 2015). In a random sample including 20,000 participants and 1,000 merchant ships, the Seafarers' International Research Centre found that two thirds of the international fleet was crewed

2 *Paper I*, section *2. Literature Review* provides an overview of shipping incidents including communication errors. It also includes references to relevant publications.

with at least two different nationalities on board, and over 20 percent of all ship types surveyed were reported to have crews of four or more nationalities (Lane, Kahveci and Sampson 2002).

Research dating back to the years 1999 and 2000 gave the proportion of non-native (L2) speakers of English as 90 percent or above (Johnson 1999; Trenkner 2000). Given that over the past twenty years or so The Philippines's and China's stake in the international shipping market has grown, the current proportion of non-native speakers of English can be assumed to have increased even further (Fan, Fei, Schriever and Fan 2017).

The Marcom project undertaken by the European Commission studied the linguistic and ethnic composition of ship crews. It found that a proportion of some 80 percent "of the world's merchant ships have become multilingual and multi ethnic in terms of crew composition" (European Commission 1999, p. 6). More recent work by Horck (2005) reports the number of multilingual crews in the merchant fleet as approximately two thirds. Sampson and Zaho provide a useful summary of the ubiquitous trend towards employing multinational ship crews:

> "Today's seafarers are commonly recruited from different world regions through networks of crewing agents and aboard modern international vessels it is common to find crews composed of men and women from several, or several dozen, countries. […] In this context it is not therefore an exaggeration to state that ships' crews have never been so nationally, culturally, or linguistically diverse"
>
> (Sampson and Zhao 2003, p. 32).

1.2. Legal framework for maritime English

Since the 18th century English has been the main language for ship communication (Molt 2006). The specific English variety spoken in the maritime domain falls into the category of English for Specific Purposes (ESP) as it is "concerned with the use of English in a restricted set of social and thematic areas chiefly for the unambiguous transfer of (technical) communication" (Gramley 2008, p. 183). In the Encyclopedia of Applied Linguistics, the language genre of *Maritime English* is defined as "an umbrella term which refers to the English language used by seafarers both at sea and in port and by individuals working in the shipping and shipbuilding industry" (Bocanegra-Valle 2013, p. 3579).

Given its importance for the shipping industry, the use of English in international shipping has been regulated by a series of legal documents.

In 1977, the International Convention for the Safety of Life at Sea (SOLAS) made the use of English compulsory if no other common language was spoken:

> "English shall be used on the bridge as the working language for bridge-to-bridge and bridge-to-shore safety communications as well as for communications on board between the pilot and bridge watchkeeping personnel, unless those directly involved in the communication speak a common language other than English."
>
> (International Maritime Organization 2014, p. 292)

In 1978, the Standard Marine Navigational Vocabulary (SMNV) was adopted by the International Maritime Organization (IMO 1978). In 1985, this document was revised, and in the year 2001 it was replaced by the Standard Marine Communication Phrases (SMCP). The aim of both the SMNV and the SMCP has been to harmonise typical speech situations on board and in radio communication in an effort to reach a better understanding in spoken communication by simplifying natural language, standardising (or *coding*) communication structures and introducing a set of message markers (Bocanegra-Valle 2011, 2010; Demydenko 2012; Franceschi 2014; Pritchard 2003; Trenkner 2002, 1996; Trenkner & Cole 2010; Winbow 2002). The use of the Standard Marine Communication Phrases was made mandatory for external communication (part A) by the IMO resolution A.918(22) (IMO 2001).

In 2010, the Manila amendments to the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW) of 1978 were adopted by the International Maritime Organization. As a minimum standard of competence for officers in charge of a navigational watch on ships of 500 gross tonnage or more, they define the following competency level to be achieved:

> "Adequate knowledge of the English language to enable the officer to use charts and other nautical publications, to understand meteorological information and messages concerning ship's safety and operation, to communicate with other ships, coast stations and VTS centres and to perform the officer's duties also with a multilingual crew, including the ability to use and understand the IMO Standard Marine Communication Phrases (IMO SMCP)."
>
> (International Maritime Organization 2010, p. 40)

By excluding the possibility to communicate in a common language other than English, the Manila amendments effectively made the use of English mandatory for the ship sizes stated above.

The IMO has also published a Model Course (3.17) on Maritime English which outlines discrete competencies for the different communicative settings on board. Unlike the competencies outlined in the STCW Convention, the IMO Model Course is not mandatory, its purpose being rather:

> "to assist maritime training institutions and their teaching staff in organizing and introducing new training courses or in enhancing, updating or supplementing existing training material where the quality and effectiveness of the training courses may thereby be improved."
>
> (International Maritime Organization 2015, p. 2)

Standardisation of maritime radio communication has also been undertaken by the International Telecommunications Union (ITU) and the International Association of Lighthouse Authorities (IALA), with "resolutions, recommendations, rules and procedures that have contributed to enhance the English language as the common language of seafaring around the world and harmonize language forms and procedures" (Bocanegra-Valle 2011, p. 37).

1.3. Communicative contexts on board sea-going ships

A sea-going ship is a complex and at the same time a very traditional work place. As Alderton puts it,

> "[h]ierarchies still persist aboard almost all vessels and the division of labour between deck and engine crews have generally been maintained. Ships retain a recognizably institutionalized form, accurately reflecting the descriptions and analyses of some of the first social scientists to look at shipboard societies."
>
> (Alderton 2005, p. 96)

The institutionalised form referred to by Alderton leads to a range of communicative contexts, depending on the ship's departments (*deck* or *engine*), hierarchical levels (*officers*, *ratings*) and performed work tasks. Particular positions aboard ship lead to divergent speech events involving different communication partners[3]. Figure 1 displays typical speech events involving navigational officers on board merchant ships.

Bridge Team Communication (BTC) is one such speech event. It refers to the spoken interaction between the nautical (*deck*) officers and/or ratings during their navigational watch on a ship's bridge (hence the term '*bridge team'*). Apart from the navigational officers involved, communication partners may also include helmsmen (*ratings*), cadets (*future nautical officers*) and marine pilots (Marine Board 1994; Weintrit 2009).

In line with the definition of English for Specific Purposes (ESP) by Gramley (2008)[4], "[t]he main purpose of bridge team communication is the exchange of information which helps the navigators increase their situational awareness in order to ensure a safe navigation" (John, Brooks, Wand and Schriever 2013, p. 242[5]).

3 Studies into Crew Resource Management have used the specific communicative contexts as categories for analysing human interaction on board ships (Chauvin, Lardjane, Morel, Clostermann, & Langard 2013; Hetherington, Flin, & Mearns, 2006; Iordanoaia 2010).

4 See chapter *1.2. Legal framework for maritime English*.

5 See chapter *10.1. Paper I*, section *4. Conclusions* for more information on the purpose of Bridge Team Communication.

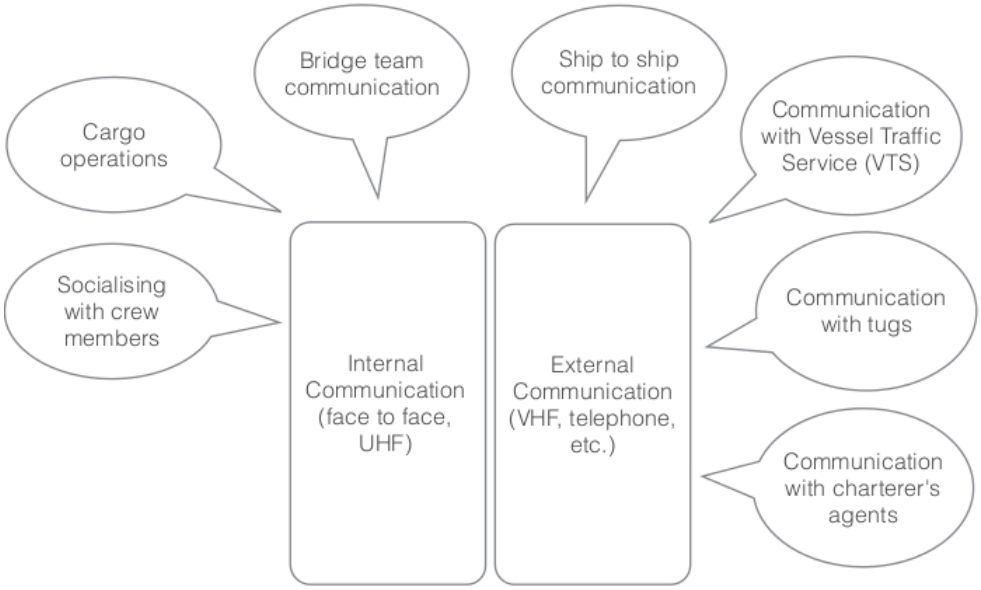

figure 1: Sample communicative context on board sea-going vessels (own illustration)

Bocanegra-Valle divides Maritime English into "five different subvarities according to the specific purpose they serve within the maritime context" (2013, p. 3580): navigation and maritime communications, commerce, law, engineering and shipbuilding. Following this categorisation, Bridge Team Communication falls into the sub-variety of "[i]nternal (intra-ship or onboard) communication" (ibid, p. 3580).

Franceschi (2014, p. 82) highlights the hybrid nature of spoken Maritime English. The author classifies it as a "standard micro-language, specifically designed to overcome or minimize miscommunication at sea" (ibid, p. 82) which "also undergoes some variation and change, possibly due to the fact that its users are speakers of other first languages, sharing the same goal but for whom intelligibility rather than correctness is key" (ibid, p. 82).

Pritchard (2003) underlines the differences between written and spoken Maritime English, and the predominant role of the spoken English register. He also relates Maritime English to field registers which he (2002, p. 3) defines as "elements of linguistic expression subjected to and imposed by the action of extralinguistic content and situational factors (field register), the speaker's role (tenor) and type /medium of communication (mode)". He uses the terms "variety" and "variant" to describe "Maritime English as a specific type and subset of general English" (ibid, p. 3). Bridge Team Communication can thus also be

classified as a micro-language which aims to reduce miscommunication at sea and as a language variant used in the specific communicative setting on a ship's bridge.

By means of radio communication, a bridge team may also be involved in other speech events. These events can include radio communication with various stations on the same ship such as the engine room, lookouts, anchor party, lifeboat stations and others. They may also include radio communication with other vessels, tugs or pilot boats and with the shore-based Vessel Traffic Services (VTS), to name but a few. On merchant ships, radio communication is usually played back through a loudspeaker system installed into the ship's navigation console, so that all bridge team members listen to radio messages although only one officer is usually in charge of speaking (Lees & Williamson 2013).

Figure 2 depicts the speech event of Bridge Team Communication on the basis of the well-known Lasswell formula which summarises the different elements of a communication process (Lasswell 1948). Following Lasswell's approach, Bridge Team Communication can be classified as a speech event in which a bridge team member (*communicator*) transmits *messages* required for the ship's navigation verbally (*medium*) to other bridge team members (*receivers*) for the purpose of information exchange (*effect*).

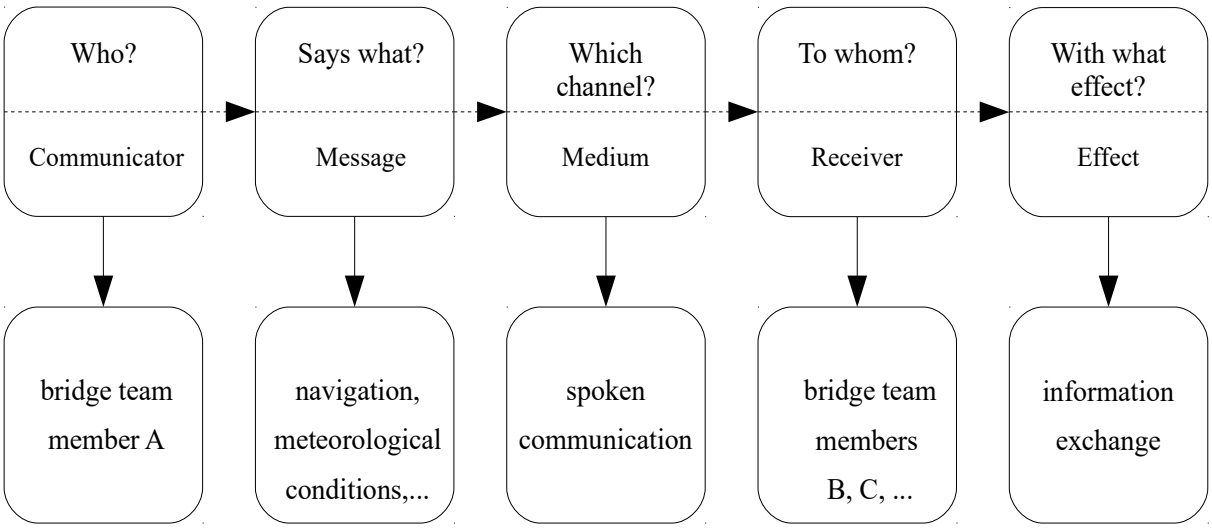

figure 2: Lasswell formula for bridge team communication
(own illustration based on the Lasswellian model)

1.4. Existing research into Bridge Team Communication

Maritime English as a language variety for specific purposes (LSP) has attracted the attention of a number of researchers. An overview of publications related to Maritime English in general is provided by Jurkovič (2015). The author categorises contributions to the International Maritime English Conference (IMEC), a subcommittee of the International Maritime Lecturers' Association (IMLA). She identifies the areas of "teaching methodology" (21% of all contributions), "intercultural and interlinguistic features" (11%) and "ME testing and assessment" (11%) (ibid, p. 194) as the most prominent in Maritime English research. Differences stated by the author between the wider LSP and the Maritime English community include an increased interest in teaching methodologies by the latter and less interest in discourse analysis in Maritime English as a genre.

Research into the empirical use of Maritime English by Strevens (1983), Strevens & Weeks (1985), Weeks (1984) and Novi (1999) has lead to major standardisation efforts in this particular LSP genre. The standardisation efforts have provided recommendations and guidelines and prescriptioncs (SeaSpeak, Standard Marine Navigational Vocabulary (SMNV), Standard Marine Communication Phrases (SMCP), VTS Guidelines) for *restricted languages* as a sub-genre of spoken Maritime English.

The Marcom (The Marcom Project 1999) project scrutinises the linguistic reality on board sea-going vessels. Its aims are to (ibid, p. 7) "assess the value of a single working language" (i.e. English) and to carry out a "linguistic analysis of ship to ship and ship to shore communications", amongst others. Recordings of radio communication segments are analysed by applying the socio-linguistic model means developed by Sinclair and Coulthard (1975). By segmenting communicative situations on board, the Marcom project has provided a framework for subsequent development of Maritime English syllabi and materials. A the same time it is acknowledged that a number of problems exist which hinder further normalisation and standardisation of Maritime English. The authors conclude by saying (The Marcom Project 1999, p. 101) that "mechanisms will have to exist to continually update a maritime English syllabus and materials".

While research into Maritime English has been conducted by a series of researchers, the number of empirical language studies into Bridge Team Communication is rather

limited, mainly due to the notorious scarcity of authentic speech data (Dževerdanović-Pejović 2013; John, Brooks and Schriever 2017[6]).

Bailey, Housley and Belcher (2006) studied authentic Bridge Team Communication by applying conversation analysis as they look specifically into confirmatory forms of talk. The authors find that "group work is oriented to achieving a mutual comprehension of the situation, with decisions flowing from a shared and interactionally generated perspective" (ibid, p. 347). In their qualitative analysis of Bridge Team Communication they highlight the importance of the temporal frame in which the communicative interaction takes place and, in general terms, they emphasize the relevance of effective maritime communication for decision-making processes on board ships.

Hontvedt and Arnseth (2013) undertake a qualitative study on the social interaction of nautical students in a ship simulator. Their analysis includes transcribed communication excerpts which are studied for the language chosen by students (English as L2 or Norwegian as L1), and the contents covered in the verbal exchange. The authors highlight the importance of role-play activities in maritime simulation and point to the fact that participants automatically chose the English language for communication when "in role" (ibid, p. 100), i.e. when they identified with their assigned rank and task, while they chose their native language "to make meta-comments" concerning the exercise (ibid, p. 100).

Research by Kataria et al. (2015) focuses on team communication between a ship's bridge and her engine control room. The researchers use semi-structured interviews in a "crew-centered design case study" (ibid, p. 175). Verbal interaction is analysed qualitatively in order to identify "the tasks undertaken by the bridge and engine room personnel in context" and to elicit "the inherent decision making required" (ibid, p. 175).

Øvergård, Nielsen, Nazir and Sorensen (2015) study participants of simulation exercises for high-speed craft. They use recorded and transcribed exercises which are analysed for their "relevance and correctness of communication" (ibid, p. 2589) by means of rubrics and propositional networks. The authors validate the adopted methodology as a possibility for assessing team performance.

6 For more information on research including authentic speech data, see chapter *10.5 Paper V,* section *I. Introduction*.

Sampson and Zhao (2003) undertake an ethnographic study by observing the social interaction and communication of multinational crews on board ship. They conclude that the knowledge of Maritime English alone is not enough for crew members to interact socially in a multilingual setting and advocate a stronger focus on general English knowledge.

In a study on pilotage operations, Sharma and Nazir (2017) identify the frequency of communication between bridge team members as an indicator for studying Distributed Situation Awareness (DSA) and bridge team performance.

Research by John, Noble and Björkroth (2016) takes on the methodology introduced in this doctoral research for analysing low-fi simulation of Bridge Team Communication. In these simulation exercises, international students of Nautical Sciences communicate in a written form via web conferencing software to assess a given navigational situation. The authors compare a number of dependent linguistic variables and find non-significant differences between communication in low-fi simulation and full-mission simulation (John, Noble & Björkroth 2013; Noble, Björkroth & John 2014). Based on the small differences between the language output in low-fi and full-mission simulation the authors see low-fi simulation as a valid tool to study spontaneous communication by future navigational officers.

Related empirical research into authentic language performance by bridge team members has mostly focused on VHF radio communication. Here, the focus is not on the verbal interaction amongst bridge team members on board the same ship but on the communication of one bridge team member speaking by radio to a shore station or to another ship. While the communicative context of radio communication is related to the face-to-face verbal interchange of Bridge Team Communication, some characteristics of VHF communication differ significantly. Maritime radio communication is much more standardised than the talk amongst bridge team members. Adherence to the Standard Marine Communication Phrases may leave room for improvement but the *restricted language* referred to by Pritchard (2002) is still clearly notable. On the other hand, radio communication is much less spontaneous than Bridge Team Communication, with clearly delineated topics to be communicated.

The most exhaustive research into VHF communication was undertaken within the

Marcom Project (The Marcom Project 1999). Based on recordings dating back to the 1990s, radio communication segments of up to 100 words each are analysed by means of the Sinclair and Coulthard socio-linguistic model (1975). The researchers are especially interested in the use of pre-scripted language structures which were found not to be applied fully by seafarers.

Brodje, Lundh, Jenvald and Dahlmann (2013) investigate "non-technical miscommunication in vessel traffic service operation" (ibid, p. 347) on the basis of transcribed audio and video recordings. The analysis is carried out by means of commercial computer software[7] to categorise different communicative settings.

Dževerdanović-Pejović (2013) adopts a qualitative discourse and intercultural approach to study authentic VHF messages[8] which have been published by marine accident investigation authorities. She studies the adoption of the prescribed SMCP by seafarers[9] and deduces that deviations from the SMCP are frequent in particular emergency situations and due to "cultural clashes" (ibid, p. 394). In her opinion, "nations or people whose collective discourse pattern is not complying or even opposing the 'Anglo-Saxon cultural script' will have difficulties in communicating in English and achieving their social position onboard" (ibid, p. 394).

Froholdt (2010) studies a telephone conversation between a vessel's shipmaster and a shore station. She adopts Wittgenstein's (1953) philosophy to carry out a qualitative analysis of the phone talk. In her research the author focuses on the display of emotions in an emergency situation. She concludes that error mitigation needs to include an "understanding [of] the social organization of a crew and the sense-making processes that take place in a context" (ibid. p. 398).

In another study, Froholdt (2015) examines the interaction of radio communication in "maritime technologically mediated interaction" (ibid, p. 468) to determine to what extent speakers use the pre-scripted language patterns of confirmatory messages. The research is carried out on audio recordings of authentic maritime radio communication.

7 The authors report to have used the MaxQDA text analysis software.

8 One of the studied transcripts has also been analysed by John, Brooks, Wand and Schriever (2013). See chapter *10.1. Paper I.*

9 At this point it is worth noting that the SMCP and the preceding SMNV were not invented by their compilers but were selected (and only slightly modified) from authentic recordings which served as the first, though modest research corpus on spoken maritime communication.

She finds that the read-back stipulations defined by the SMCP are generally deployed by seafarers.

Kataria (2011, 2015) undertakes an ethnographic research into radio communication with Vessel Traffic Services in India to explore the "social order and […] the local practical achievement of traffic coordination and channel navigation in restricted waters" (Kataria 2015, p. *XIX*). Her research pursues an ethnomethodological approach to analyse "observations, interviews and the real-time naturally occurring interaction on the port VHF radio" (ibid, p. 6).

Pyne and Koester (2005) study maritime accident reports by applying an investigation taxonomy commonly used in the aviation industry. The authors apply the ADREP 2000 taxonomy to classify miscommunication in multicultural crews and conclude that

> "[t]he multinational crew must interact and communicate in a common language to maintain "social harmony' in an off duty context and in their everyday 'teamwork' to ensure effective day to day operation. The most commonly recognised failure occurs with the level of understanding of English between ship to ship and/or ship to shore under conditions of restricted manoeuvrability, or when under critically congested circumstances where little time or space can be afforded for mistakes to be made."
>
> (Pyne and Koester, p. 206)

Further review of relevant research outside a maritime communicative context such as the aviation industry, nuclear power plants and other areas in which communication is a decisive factor are included in the respective publications of this dissertation[10].

10 See chapter *10. Appended papers*.

1.5. Research gaps

While all researchers mentioned in chapter 1.4 agree on the importance of effective maritime communication in an intercultural and multilingual setting, nearly all presented methods apply qualitative, case-based methods or taxonomies to make judgements on the quality of the studied communication fragments. By adopting a qualitative approach, the differentiation between effective and ineffective maritime communication relies on an expert judgement made by the researchers who classify or evaluate verbal interaction by means of introspection (Faerch and Kasper 1987, Butler 2013). The publications' findings are highly relevant and most suitable to make inferences on Bridge Team and VHF communication in general as they identify areas for improvement which can be taken into account when developing training courses or curricula for Maritime Education and Training institutions. However, none of the research presented above aims to define quantitative measures to compare empirical language performance against a standard or benchmark value[11]. A quantitative approach offers the advantage that once the method has been defined and validated it can subsequently be applied without the need of an introspective expert judgement. Quantitative research into communication has the potential to gauge observations of communicative performance by delivering numerical values and statistical distributions which can be compared to values and distributions calculated for other individuals, sociolinguistic groups or text genres, amongst others.

The identified lack of quantitative methods for assessing language performance by bridge team members leads to the first and principal research gap ($RG_{1)}$:

$\Rightarrow RG_1$ No quantitative methods have been defined and validated to study naturalistic language performance by bridge team members.

Most research into Bridge Team or VHF radio communication is based on individual cases. Inferences are made on small sample sizes (often one single case). Again, an introspective expert judgement based on the researchers' personal and professional experience is required to categorise the studied observations as *appropriate* and the

11 With the exception of research into low-fi simulation exercises by John, Noble and Björkroth (2013, 2014, 2016) which has adopted the methodology presented in this dissertation.

communicative behaviour as *idiosyncratic* for bridge teams. To the contrary, a Corpus Linguistics approach is suitable to identify idiosyncratic language patterns on the basis of a much bigger sample size involving a collection of Bridge Team Communication by a number of speakers. The lack of a Corpus Linguistics approach to investigate Bridge Team Communication leads to the second research gap (RG_2):

⇒ RG_2 No Corpus Linguistics approach has been adopted for studying naturalistic, inter-personal communication by bridge team members.

Corpus Linguistics studies empirical language use beyond individual cases and it seeks to identify patterns which are deemed typical for specific sociolinguistic groups, speech communities, genres, etc. When studying individual cases, the research publications cited in chapter *1.4. Existing research into Bridge Team Communication* do not compare native (L1) and non-native (L2) speakers of English in a methodological and comprehensive manner. It is taken for granted that native speakers provide the linguistic framework at which non-native speakers of English should aim and for which they are trained. However, no empirical data is used to contrast L2 speakers to authentic language performance provided by their native counterparts while they carry out identical navigational tasks. This lack of comparative analyses leads to research gap number 3 (RG_3):

⇒ RG_3 No empirical Corpus Linguistics research has been conducted on differences and similarities in language performance by native and non-native speakers of English engaged in identical navigational tasks.

Finally, an assessment of authentic language production in a complex professional environment needs to consider linguistic factors which potentially lead to communicative disruptions and miscommunication. The case-based, introspective approach adopted by existing research into maritime communication does not include any attempt to quantify observations of ineffective communication which eventually lead to an increased risk to crew members, shipping companies and the environment. This lack of a quantitative risk analysis of miscommunication leads to the final research gap (RG_4):

⇒ RG_4 No quantitative risk analysis has been carried on the linguistic structures of native and non-native speakers of English engaged in identical navigational tasks.

1.6. Research questions and hypotheses

In the following, principal and subordinate research questions are formulated which aim to fill the research gaps identified in the previous chapter. References are included to relate these questions to the pertinent research publications.

RG_1 has identified a lack of quantitative methods to assess naturalistic language performance by bridge team members. Research question 1 (RQ_1) targets this particular gap:

⇒ RQ_1 How can idiosyncratic linguistic patterns of Bridge Team Communication be modelled by means of quantitative Corpus Linguistics methods?

RQ_1 is the principal research question and is thus studied in all papers submitted in fulfilment of this doctoral research. Table 1 lists subordinate research questions to RQ_1 which are investigated in the respective papers.

Research question	Subordinate research questions	Research papers
⇒ RQ_1	$RQ_{1.1}$: Can the effectiveness of Bridge Team Communication be assessed by means of a specific lexical index measuring the exchange of maritime information?	I, III, IV
	$RQ_{1.2}$: Can the grammar diversity produced by different speakers be compared by means of a special part-of-speech diversity index?	II, III, IV

table 1: research question 1 and subordinate research questions

RG_2 relates to the lack of a Corpus Linguistics approach to study naturalistic, inter-personal communication by bridge team members. This gap is targeted by research question 2 (RQ_2):

⇒ RQ_2 Can quantitative Corpus Linguistic methods identify suitable patterns to assess the language performance and thus the effectiveness of Bridge Team Communication?

RQ_2 refers to the application of quantitative tools to a specific text corpus. In quantitative research, the applied methods are expected to produce statistically significant results which have the potential to differentiate between the language performance observed in individual bridge team members or teams, or between sociolinguistic groups. While this chapter defines general, overarching research questions, in the stated research papers these questions are broken down further into several null and alternative hypotheses ($H_0 \neq H_a$)[12]. The magnitude of the computed effects is further calculated by means of appropriate statistical methods so that the observed effects provide numerical results which can be compared and ranked. Table 2 lists subordinate research questions to RQ_2 which are studied in the respective papers.

Research question	Subordinate research questions	Research paper
⇒ RQ_2	$RQ_{2.1}$: To what extent do the speech patterns of bridge team communication by non-native speakers of English in full-mission simulation differ lexically and grammatically from other, non-nautical communication?	III
	$RQ_{2.2}$: Can the exchange of technical information by bridge team members be assessed by means of Corpus Linguistics methods?	I, II III, IV
	$RQ_{2.3}$: Can nautical idiomaticity be assessed in authentic, inter-personal communication by bridge team members?	I, III, IV

table 2: research question 2 and subordinate research questions

12 See chapters *10.3. Paper III* and *10.4 Paper IV*.

This gap is targeted by research question 3 (RQ_3) and its subordinate research questions listed in table 3.

⇒ RQ_3 What differences can be observed in the linguistic patterns produced by bridge teams composed either of native speakers or of non-native speakers of English while performing identical navigational tasks?

By following the methodology outlined in research papers I and II and replicating the approach adopted in paper III, the following subordinate research questions directly relate to $RQ_{1.1}$ and $RQ_{1.2}$ as well as to $RQ_{2.1}$, $RQ_{2.2}$ and $RQ_{2.3}$.

Research question	Subordinate research questions	Research paper
⇒ RQ_3	$RQ_{3.1}$: To what extent do the speech patterns of bridge team communication by non-native speakers of English differ lexically and grammatically from those produced by native speakers performing identical navigational tasks in full-mission simulation?	IV, V
	$RQ_{3.2}$: In how far does the exchange of technical information by native speakers of English differ from that of non-native speakers?	IV, V
	$RQ_{3.3}$: Can a difference in nautical idiomaticity by native and non-native speakers of English be observed in authentic, inter-personal communication by bridge team members?	IV, V

table 3: research question 3 and subordinate research questions

Research questions RQ_1 to RQ_3 aim at determining and validating a quantitative methodology capable of identifying significant differences between individuals, bridge teams or sociolinguistic groups (namely native and non-native speakers of English). The definition of quantitative methods for assessing bridge team communication should to some extent offer the possibility to eliminate the need of a case-based, introspective approach in assessing effectiveness of Bridge Team Communication.

Research question 4 (RQ_4) seeks to identify potential risks of communicative disruptions and miscommunication by bridge team members. This risk identification complements the other research questions by estimating communicative competence of teams or sociolinguistic groups.

⇒ RQ_4 How can quantitative Corpus Linguistics methods assist in identifying risks of communicative disruptions and miscommunication by bridge team members?

RQ_4 considers dialogues beyond an utterance level and builds on the speech act theory by Austin (1962) and Searle (1969)[13]. Table 4 lists the pertinent subordinate research questions dealing with risk analysis in communication patterns:

Research question	Subordinate research questions	Research paper
⇒ RQ_4	$RQ_{4.1}$: How can risks of miscommunication by bridge team members be identified by means of quantitative Corpus Linguistics methods?	V
	$RQ_{4.2}$: How can differences in the locutionary and illocutionary meaning be used to assess a potential risk of miscommunication by bridge team members?	V
	$RQ_{4.3}$: How can lexical structures in perlocutionary speech acts be used to make inferences on an actual miscommunication by bridge team members?	V

table 4: research question 4 and subordinate research questions

13 See chapter *10.5 Paper V* for an introduction into Austin and Searle's Speech Act theory.

1.7. Outline of the thesis

The first chapter concludes by providing an overview of the following chapters and the adopted approach to answer the research questions stated above.

Chapter *2. Theoretical framework* highlights the importance of professional team communication by referring to the theoretical constructs of situational awareness and shared mental models. It also introduces the selected communication model for studying Bridge Team Communication from an information exchange perspective. Further, a definition of Discourse Analysis is provided as a means to investigate spontaneous verbal communication. The chapter concludes by applying the theoretical framework to the specific discourse community of bridge team members.

Chapter *3. Research Methods and Materials* defines the methodological research framework by introducing essential concepts of Applied Linguistics and Corpus Linguistics. This chapter also refers to the notions of quantitative content analysis and frequentist inferencing as two major analytical tools employed in quantitative research. Reference is also made to the cross-sectional and mixed-method approaches followed in collecting and analysing naturalistic spoken discourse. The chapter concludes by categorising the adopted research methods into an applicable research epistemology.

Chapter *4. Data collection and data processing* explains how the naturalistic spoken discourse data were recorded and transcribed. It continues by outlining how the primary data were processed in order to build a special spoken text corpus of Bridge Team Communication. This chapter also reflects on the corpus' representativeness including aspects related to the corpus' size, to its lexical closure and dispersion figures. Furthermore, the corpus mark-up strategy is introduced including the employed part-of-speech, key word and pragmatic tagging.

Chapter *5. Data analysis* describes the used methods for analysing the developed Bridge Team Communication corpus. It details the application of quantitative content analysis and it refers to the use of frequencies and proportions to make inferences regarding lexical and grammatical linguistic patterns. The chapter also refers to the statistical tools which have been employed to analyse the linguistic and communication structures by the sampled bridge team members.

Chapter *6. Measures of research quality and robustness* looks into measures of research soundness. The reliability of adopted research methods is analysed by studying both the research instruments and its results. Further, the validity of the research tools is studied internally and externally, and reflections are made as to the replicability and generalisability of the results to other discourse communities.

Chapter *7. Published results* summarises the main findings of the research papers. This chapter also details their inter-relationship and describes their contribution towards answering the research questions outlined above. The chapter concludes by presenting the overall results of the research papers, namely the application of a probabilistic linguistic profile for assessing naturalistic spoken discourse.

Chapter *8. Discussion and conclusions* closes the research thesis by providing a discussion and conclusion on the implications and on the constraints and limitations of the adopted research methodology. The chapter also includes remarks on possible future research in this particular domain and finishes with final observations.

Chapter *9. References* lists the used research literature while chapter *10. Appended papers* includes the research papers which constitute the core part of this thesis.

Finally, chapter *11. Appendices* includes the forms used to invite participants to this research, give their consent and provide their demographic data.

2. Theoretical framework

Chapter two introduces the theoretical foundations on which this research builds. It refers to the notions of situation awareness and shared mental models and it highlights the importance of interpersonal communication for these two constructs. It also introduces the selected model for describing verbal communication and it outlines key terms related to research of spoken discourse. The chapter concludes by blending the introduced cognitive and linguistic concepts, and outlining their importance for bridge team communication.

2.1. Importance of professional team communication

Much of today's work (including the shipping industry) is carried out in dynamic systems which "rely on the performance of teams rather than on individuals" (Garbis & Artmann 1998, p. 151). The complexity of dynamic socio-technical systems demands highly specialised professional teams which need to make decisions and adopt appropriate co-operative strategies. Decisive information required by team members to reach informed decisions is primarily exchanged by means of spoken discourse. If the exchange of relevant information is asynchronous or imbalanced, it may cause a divergent mental representation of a given situation by the different members of a team. The primary role of professional team communication is thus the creation of a *shared mental model* (Orasanu 1990), i.e. a common understanding of the tasks and challenges at hand. Mental models can be seen as "organized knowledge structures that allow individuals to interact with their environment" (Mathieu, Heffner, Goodwin, Salas & Cannon-Bowers 2000, p. 274). They "provide a context in which communication can be interpreted, and a basis for predicting the behavior and needs of other members" (Cannon-Bowers, Salas & Converse 1993, p. 229). The theoretical construct of shared mental models has been studied in different professional domains including the aviation (Stout, Cannon-Bowers, Salas & Milanovich 1999) and health care industries (Haig, Sutton & Whittington 2006).

Shared mental models are closely related to the construct of a common *situation awareness (SA)*. Endsley summarises the importance of situation awareness as follows:

> "In dynamic environments, many decisions are required across a fairly narrow space of time, and tasks are dependent on an ongoing, up-to-date analysis of the environment. Because the state of the environment is constantly changing, often in complex ways, a major portion of the operator's job becomes that of obtaining and maintaining good SA".
>
> (Endsley 1995, p. 33)

The concept of situation awareness has been studied extensively in aviation (Adams, Tenney & Pew, 1995; Endsley 1997; Endsley & Jones 2013; Jenson 1997).

Individual situation awareness refers to the integration of environmental information and combining it with "previous knowledge to form a coherent mental picture" (Dominguez 1994, p. 11). The resulting mental picture of an individual may be seen as a "generative process of knowledge creation and informed action taking" (Smith & Hancock 1995, p. 142). By contrasting a developed mental picture with the real development of a given situation, a "perceptual cycle" (Smith & Hancock 1995; Niesser 1976) is generated as a holistic process of mental processes and products (Salmon et al 2008, p. 303).

While the construct of individual situation awareness refers to the perceptual cycle of one person (or team member), the construct of *team* situation awareness refers to the "common picture" (Salmon et al 2008, p. 308) of all team members as a multi-dimensional "shared understanding" (ibid, p. 309).

The two theoretical constructs of *shared mental models* and *team situation awareness* both rely on effective communication amongst team members (Bolstad & Endsley 2000; Endsley 1995; Entin & Entin 2000; Salas et al 1995) to develop a common understanding of a given professional situation and measures to be taken in order to deal with existing and future tasks and challenges in a complex socio-technical work environment[14].

14 See chapter *10.4 Paper IV*, section *2. Definition and measurement of cognitive load* for further information and references on psychological constructs related to the socio-technical environment on board sea-going ships.

2.2. Adopted communication model

Defining the concept of a common term like *communication* has been attempted by quite a number of researchers. As early as in the year 1976, Dance & Larson had already spotted 126 different definitions of *communication* (Frey, Botan and Kreps 2000). Pearce denotes a

> "difference in the connotations of communication depending on whether the emphasis is on that which is made *common* (shared meanings, cultural symbols, traditions, common ground, understanding) or on the process of *making* things common (the transmission of messages from place to place; the languages in which things are framed; the patterns of actions in which they occur; the things that people actually do and say to each other)."
>
> (Pearce 1995, p. 7).

While the former approach describes a "meaning-based or constitutive perspective", the latter focuses on communication from an "information exchange perspective" (Mokros and Deetz 1996, p. 32).

As this research concentrates on a communicative setting in which bridge team members share navigational information, it follows the information exchange perspective outlined by Pearce and Mokros & Deetz. For this purpose, Gerbner's simple definition of communication as a "social interaction through messages" (1967, p. 41) is best suited as it takes into account the *sender-receiver* interaction and the corresponding *message* or information transmitted.

In social sciences, a model may be considered "a consciously simplified description in graphic form of a piece of reality. A model seeks to show the main elements of any structure or process and the relationships between these elements" (McQuail and Windahl 1993, p. 2). Deutsch (1966) hints at the organising function of models as well as their explanatory nature to disambiguate complex information.

In communication models dealing with information exchange the main elements of *sender/transmitter – message – receiver* constitute the very core of the communicative structure. One of the first models to include these core elements is the well-known Transmission Model developed by Shannon and Weaver (1949). Johnson and Klare (1961) highlight the model's importance in line with Shannon's mathematical approach

towards communication. Criticism on the Transmission Model dwells on its linear structure which does not include any feedback channel (Dance 1967; DeFleur 1970; Schramm 1954).

A model that takes into account the reciprocity of human communication was developed by Osgood and Schramm (1954, see figure 3). The circular Osgood & Schramm model does not only include the *sender*, *receiver* and the respective *messages*, it also incorporates the relevant aspects of *encoding*, *interpreting* and *decoding* the transmitted messages. From an information exchange perspective, communication can be deemed to have been successful when messages are encoded correctly by the sender *and* when they are decoded and interpreted correctly by the receiver. Conversely, miscommunication takes place when a temporary or permanent disruption of the information exchange occurs by coding or interpreting messages incorrectly[15].

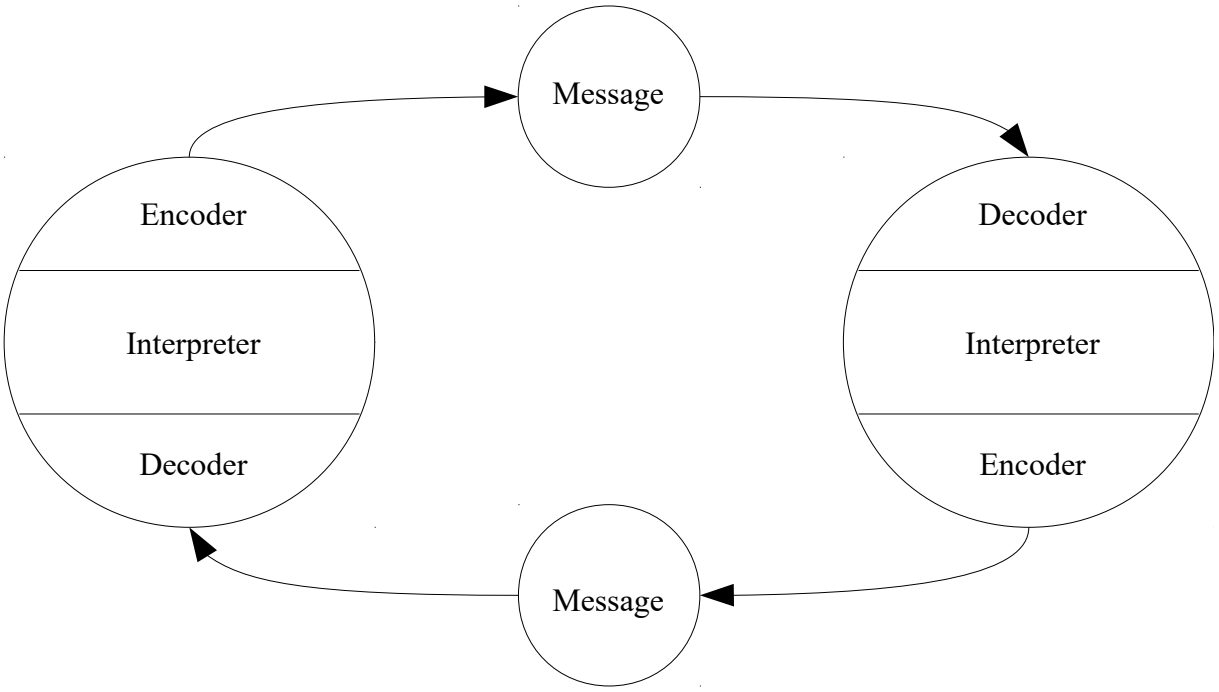

figure 3: Osgood & Schramm model of communication
(own illustration based on Schramm)

Like other models in social sciences the Osgood & Schramm model consciously simplifies a piece of reality for the sake of clarity. A shortcoming highlighted by Dance (1967, p. 295) refers to the model's circularity as it "suggests that communication comes back, full circle, to exactly the point from which it started". Hence, no progress in the communicative process is expressed.

15 See chapter *10.5 Paper V* for an analysis on errors related to coding and decoding messages.

Another aspect which is not explicitly included in the Osgood & Schramm model is the possible existence of noise or other external factors related to the transmission quality which might affect the receiver's decoding capabilities negatively. However, in can be argued that these external factors can possibly be considered to have been integrated into the decoding phase of the communicative process.

With regards to providing a qualitative model for Bridge Team Communication, the Osgood & Schramm model is quite suitable for categorising speech acts into different stages. A bridge team member encodes a message and transmits it verbally to another bridge team member. The latter decodes the message and interprets it. Upon interpreting the message he or she will provide feedback by encoding a response and transmitting a reply message back to the original sender. In shipping, a communicative process providing feedback loops is referred to as *closed loop* communication whereby the sender's original message is acknowledged by the receiver by repeating its most important parts or by repeating the message entirely (Bowers, Jentsch, Salas and Braun 1998; Brodje, Lundh, Jenvald and Dahlman 2013; Porathe, Eklund and Goransson 2014).

Following the definitions summarised in this chapter, Bridge Team Communication is studied in this research as an information exchange between bridge team members. For this purpose, the spontaneous spoken exchange of navigational information is categorised into messages which are coded and sent as utterances by one member of a bridge team and received and decoded by another member of the same team.

2.3. Analysis of spoken discourse

Discourse Analysis (DA) is a widely used methodology which is not only employed by linguists but which is also common to other social sciences such as psychology, sociology and educational studies (Hymes 1986, Swales 1986, 1990). The term *discourse analysis* itself was coined by Harris (1952) in his research into "language beyond the level of the sentence and the relationship between linguistic and non-linguistic behaviour" (Paltridge 2006, p. 2). Cameron (2001, p. 11) explains the expression "language above the sentence" as a quest to identify "patterns (structure, organization) in units which are larger, more extended, than one sentence". She suggests the following definition: "language in use: language used to do something and mean something, language produced and interpreted in a real-world context" (ibid, p. 13). She further describes the holistic nature of Discourse Analysis as a

> "method for doing social research; it is a body of empirical knowledge about how talk and text are organized; it is the home of various theories about the nature and workings of human communication, and also of theories about the construction and reproduction of social reality. It is both about language and about life."
>
> (Cameron 2011, p. 17).

Discourse Analysis is a most adequate tool for analysing communication processes as depicted by the Osgood & Schramm model. A circular speech event including one or more feedback loops needs to be studied *above the sentence* level, and verbal social interaction by bridge team members needs to be investigated by "taking account of the social and cultural setting in which the speaking or writing occurs, speakers' and writers' relationships with each other, and the community's norms, values and expectations for the kind of interaction, or speech event" (Paltridge 2006, p. 6).

The *language in use* aspect hints at another important facet of discourse analysis. By concentrating on actual language use, discourse analysis does not exclusively deal with *linguistic* competence but also with *communicative* competence (Hymes 1972). While the former term refers to the grammatically correct use of language, the latter considers the actual, naturalistic use of language in a specific discursive context.

2.4. Application of the theoretical construct to Bridge Team Communication

Navigating a sea-going ship is a co-operative task carried out by bridge team members. Individual members need to share a common mental representation of the navigational situation in order to safeguard a high situational awareness level by all team members. To achieve such a shared mental model, bridge team members need to verbally communicate their own assessment of a given situation with other members of the team. This information exchange process can be categorised into discrete elements by means of the Osgood & Schramm communication model. The model differentiates between senders and receivers of messages, and it takes into account the reciprocity of the communication process.

An assessment of naturalistic communication processes can be undertaken by means of Discourse Analysis. This technique is suitable to study the language in use above the sentence level and in a social and cultural setting.

3. Research Methods and Materials

This chapter starts off by describing the adopted research strategy and the pertinent methodological assumptions. Subsequently, the adopted research paradigm and its inherent approach is outlined including aspects related to Applied Linguistics and Corpus Linguistics as well as the adopted cross-sectional and mixed-methods approach. Reference is also made to the Sociolinguistics, Psycholinguistics and Corpus Pragmatics methods employed in the research papers. The chapter closes by classifying this doctoral research into distinct research categories.

3.1. Methodological framework

Chapter *2. Theoretical framework* introduced the underlying theories and constructs of this research. For the study of language performance based on empirical spoken discourse data, these qualitative concepts are investigated and validated by implementing the quantitative methodology outlined in this chapter.

Following a classification by Dörnyei (2007), quantitative research includes six characteristic features: It focuses on *numbers*. This means, that the linguistic aspects to

be studied are classified and consequently quantified, thus providing numerical values. The classification is carried out as an *a priori categorisation* which means that prior to the data collection, the measurement instruments are defined and subsequently applied to the investigated discourse data. Quantitative research is carried out on *variables rather than on cases*. The focus lies on "common features of groups of people" (ibid, p. 33) rather than on individuals. The numerical values provided by the measurement instruments are analysed *statistically* to identify patterns and structures in the scrutinised language variety. Quantitative analyses employ *standardised procedures to assess objective reality* in an effort to reduce the dependence on "idiosyncratic human variability and bias" (ibid, p. 34). Finally, inferences are made on samples and generalised to determine *universal laws* of human interactive communication in the particular sociolinguistic domain of Bridge Team Communication.

Apart from this general description of quantitative research, the methodological framework adheres to a number of different approaches which are detailed in the following.

3.2. Applied Linguistics approach

The research activities and findings which constitute this dissertation take place in the framework of *Applied Linguistics*, in line with the definitions by Davies and Elder (2008, p. 1) who see this specific linguistics field as "concerned with solving or at least ameliorating social problems involving language". In a similar fashion, Phakiti and Paltridge (2015, p. 6) state that "[a]pplied linguistics (AL) provides the theoretical and descriptive foundations for the investigation and solution of language-related problems". They continue by saying that:

> "[a]pplied linguistics (AL) is an interdisciplinary field of study that aims to understand the multifaceted roles and nature of language use and/or language problems in social contexts (see e.g. Berns & Matsuda 2006; Cook 2003; Davies & Elder 2004; Hall, Smith & Wicaksono 2011; McCarthy 2001; Pawlak & Aronin 2014; Pennycook 2001; Schmitt 2002 for detailed discussions). AL researchers are interested, for example, in understanding how language is used or learnt as well as what problems or difficulties people face when using language to communicate in a variety of situations and contexts. To achieve this, applied linguistics researchers draw on a range of theories and research

> methodologies not only from linguistics, but also from other disciplines such as education, psychology and sociology."
>
> (Phakiti & Paltridge 2015, p. 5)

Research in Applied Linguistics is the "organized, systematic search for answers to the questions we ask" (Hatch and Lazaraton 1991, p. 1), including an "inquiry consisting of three elements of components: (1) a question, problem or hypothesis; (2) data; and (3) analysis; and interpretation of data" (Nunan 1992, p. 3.) for the "theoretical and empirical investigation of real-world problems in which language is a central issue" (Brumfit 1997, p. 93).

The focus of Applied Linguistics on the analysis and interpretation of empirical, primary data is especially visible in its "concern with professional activities which aim to solve real-world language-based problems" (Phakiti and Paltridge 2015, p. 6) in an inter-disciplinary manner, or, as Schmitt and Celce-Murcia (2002, p. 1) describe it, "in order to achieve some purpose or solve some problem in the real world".

Instead of providing new theoretical frameworks, Applied Linguistics uses "knowledge or theories from basic research to address a problem by systematically applying them through activities with a group of individuals and observing how they work to, for example, enhance learning or improve a process" (Phakiti and Paltridge 2015, p. 11). Kaplan (2002, p. 514) underlines this approach by saying that Applied Linguistics "are likely to move toward the analysis of new data, rather than continue to argue new theory".

This doctoral research follows the research paradigm of Applied Linguistics for providing a quantitative model of Bridge Team Communication based on empirical data of naturalistic speech patterns. The verbal interaction observed in bridge team simulation exercises is analysed and profiled[16] based on the defined methods[17] to identify ineffective communication and miscommunication[18]. By identifying miscommunication in a quantitative manner this research aims to contribute to the solution of the real-world language-based problems referred to above by Phakiti and Paltridge and make bridge

16 For more information on linguistics profiling, see chapters *10.3. Paper III* and *10.4 Paper IV*.

17 For more information on the defined methods, see chapters *10.1. Paper I* and *10.2. Paper II*.

18 Findings of ineffective communication and miscommunication are detailed in chapter *10.5 Paper V*.

team communication more effective and hence safer.

3.3. Corpus linguistics approach

In this research, objective measurements of authentic language use are pursued by applying Corpus Linguistics techniques. These techniques include corpus processes described by Baker (2006, p. 1) as "computational procedures which manipulate […] data in various ways [...] in order to uncover linguistic patterns which can enable us to make sense of the ways that language is used in the construction of discourses (or ways of constructing reality)". Stubbs (2008, p. 106) sees the advantage of text corpora in providing "observable evidence about language use, which leads to new descriptions, which in turn are embodied in dictionaries, grammars and teaching materials". Corpus Linguistics deals with the analysis of language patterns and with "findings about recurrent lexico-grammatical units of meaning which have implications for both theoretical and applied linguistics" (ibid, p. 106). According to Biber, Conrad and Reppen (1998, p. 4), linguistic patterns "represent quantitative relations, measuring the extent to which features and variants are associated with contextual factors".

Criticism on Corpus Linguistics techniques includes Chomsky's famous quote: "if you sit and think for a few minutes, you're just flooded with relevant data" (1984, p. 44). The rationalist approach advocated by Chomsky suggests introspection as an adequate way of eliciting linguistic data. McEnery and Wilson (2001) accept Chomsky's rationalist approach for studies into language *competence*. However, most Corpus Linguistics research does not deal with language competence but with language *performance*, i.e. the naturalistic use of language as referred to by Baker and Stubbs above. The dichotomy of studies on language competence as "tacit internalised knowledge" (Baker, Hardie and McEnery 2006, p. 39) and on language use as "behaviour in real life" (ibid) has eventually been recognised by Chomsky (1965, 1988) who also uses the terms *externalised language* (E-language) and *internalised language* (I-language).

Leech (1992, p. 107) summarises the research approach adopted by Corpus Linguistics as follows:

„(1) Focus on linguistic performance, rather than competence

(2) Focus on linguistics description, rather than linguistic universals

(3) Focus on quantitative, as well as qualitative models of language

(4) Focus on a more empiricist, rather than rationalist view of scientific inquiry"

Tognini-Bonelli (2001) makes a distinction between *corpus-based* and *corpus-driven* research. While the former uses text corpora as a "repository of examples to back pre-existing theories or a probabilistic extension to an already well defined system" (ibid, p. 84), the latter expresses "the commitment of the linguist […] to the integrity of the data as a whole, and descriptions aim to be comprehensive with respect to corpus evidence" (ibid, p. 84). Instead of being limited to providing evidence for constructs often being inferred by introspection, corpus-driven research aims to

> "reflect directly […] the evidence provided by the corpus. Indeed, many of the statements are of a kind that are not usually accessible by any other means than the inspection of corpus evidence. Examples are taken verbatim, in other words they are not adjusted in any way to fit the predefined categories of the analyst; recurrent patterns and frequency distributions are expected to form the basic evidence for linguistic categories; the absence of a pattern is considered potentially meaningful."
>
> (Tognini-Bonelli 2001, p. 84).

This research carried out on Bridge Team Communication studies the language *performance* of the audio-recorded bridge teams quantitatively. It analyses frequency data such as the proportions of the studied variables to the overall number of items under scrutiny. By doing so in a *corpus-driven* fashion, it "aims to derive linguistics categories systematically from the recurrent patterns and the frequency distributions that emerge from language in context" (Tognini-Bonelli 2001, p. 89). In accordance with Baker and Ellece, the researching student

> "approaches data with an 'open mind' and allows whatever emerges as interesting, salient

or frequent to 'drive' the analysis along. The data then direct the analyst to choose certain features or adopt a particular analytical framework."

(Baker & Ellece 2011, p. 29)

3.4. Cross-sectional approach

The research presented in this dissertation is *cross-sectional* as it involves data "from one or more cohorts (a person, group of people) at a single point in time or within a short period of time" (Phakiti and Paltridge 2015, p. 12). The adopted *synchronic* approach aims to provide a "contemporary language description" (Liddicoat and Curnow 2008, p. 25) by portraying the use of language "as it is at a particular moment in time" (ibid).

The adopted cross-sectional approach infers results from the sampled verbal communication by native and non-native speakers of English which was recorded in the years 2013 and 2014[19]. On the basis of the transcribed synchronic language data, a series of corpus-driven analyses is conducted on differences

- between maritime (bridge team) and non-maritime communication[20],
- between native and non-native speakers of English[21], and
- on risks of miscommunication caused by communicative disconnects and breakdowns[22].

3.5. Mixed-methods approach

Where appropriate, inferences on the statistical findings are made by *triangulating* the adopted quantitative approach with interpretative research methods such as qualitative Discourse Analysis (McNeill 1990; Newby 1977). Layder (1993) sees the advantage in triangulation in an improved validity check of hypotheses, anchoring findings in better interpretations and explanations and a higher flexibility in responding to unexpected occurrences. Baker and Egbert (2016, p. 3) find that "most contemporary corpus

19 For more details on the sampling, see chapter *10.3. Paper III,* section *2.2 Bridge team transcript*

20 See chapter *10.3. Paper III* for a detailed analysis.

21 See chapter *10.4 Paper IV* for a detailed analysis.

22 See chapter *10.5 Paper V* for a detailed analysis.

linguists employ triangulation to an extent in their own research by, for example, using different techniques on their corpora".

Table 5 provides an overview of the applied research methods following Brown's (2004, p. 496) "Standards of research soundness continua for primary research". The elements which apply in this doctoral research are marked in bold characters.

Primary research		
↓		
Interpretative research	**Survey research**	**Statistical research**
↓	↓	↓
• Case studies	• Interviews	• **Descriptive**
• Introspection	• **Questionnaires**	• **Exploratory**
• **Discourse analysis**		• Quasi-experimental
• **Interactional analysis**		• Experimental
• Classroom observations		

table 5: Standards of research soundness continua for primary research (adapted from: Brown 2004)

In the present research on linguistic structures of Bridge Team Communication a combination of quantitative and qualitative data has been used to interpret statistical findings of quantitative analyses in an "explanatory sequential design" (Creswell 2014, p. 6) by which data is firstly analysed by means of quantitative methods and subsequently employing qualitative methods to explain the observations in a more detailed manner.

3.6. Epistemology

The outlined application of multiple strategies for gathering and analysing data follows a post-constructivist epistemology (LeCompte and Schensul 2010; Lincoln, Lynham and Guba 2011). The post-constructivist mindset aims to guarantee the highest possible degree of objectivity within a critical realist ontology but concedes that reality is often more complex than the models offered by researchers (Denzin 1994; Guba and Lincoln 1994; Phakiti and Paltridge 2005). To provide for a multi-faceted view on this complex reality, a variety of methods is used, drawn from the domains of

Sociolinguistics[23], Psycholinguistics[24] and Pragmatics[25]. By studying a given communicative setting from different angles (i.e. linguistic branches) the complex reality of language is critically analysed to achieve an elevated degree of objectivity.

In a nutshell, this research can be categorised as follows:

- it takes place in the framework of Applied Linguistics and includes the areas of Sociolinguistics, Psycholinguistics and Pragmatics,
- it applies knowledge and theories of existing basic research to provide solutions to real-world problems based on communication,
- it is based on empirical, primary data analysed in cross-sectional research on synchronic language use,
- it applies quantitative methods and Corpus Linguistics techniques which are triangulated with interpretative methods, and
- it adopts a post-constructivist epistemology and a critical realist ontology.

4. Data collection and data processing

As stated in chapter *3.1. Methodological framework*, this research includes quantitative discourse analyses of naturally occurring speech patterns observed in Bridge Team Communication exercises. To date, very little transcribed data of empirical bridge team interaction is available in the public domain, as none of the large reference corpora (e.g. the British National Corpus, the Brown Corpus family, etc.) contains any maritime communication. The few existing corpus-based studies on maritime communication[26] also do not reveal their primary data. Even if a very small number of accident investigation reports include transcripts of verbatim data (Marine Accident Investigation Branch 2010, National Transportation Safety Board 2008a, 2008b, 2009), these samples

23 See chapter *10.3. Paper III*.

24 See chapter *10.4 Paper IV*.

25 See chapter *10.5 Paper V*.

26 See chapter *1.4. Existing research into Bridge Team Communication*.

are insufficient to conduct comprehensive linguistic analyses due to their limited size or to their lack of contextualisation (Hardt-Mautner 1995; Leech 1991; Partington 2003).

4.1. Collection of primary data (sampling)

To overcome this shortage and to provide for a "study of language based on examples of real life language use" (McEnery and Wilson 1996, p.1), a specialised text corpus has been built and which aims to deliver a "particular representative function" (Leech 1991, p. 11) of the linguistic genre studied. A purpose-built text corpus on Bridge Team Communication offers the additional advantage of having full control over the sampling process while minimising de-contextualisation of the collected data (Hardt-Mautner 1995; Partington 2003), or as Baker (2007, p. 31) puts it: "reference corpora may not contain enough of the text types you are interested in examining or may not have enough references to the subject(s) you want to investigate".

In order to achieve a well-balanced composition of the Bridge Team Communication text corpus, preliminary analyses of suitable simulation exercises were carried out for an assessment of both the technical aspects of audio-recording participants (i.e. the sound quality and speech intelligibility) and the nautical scenarios to be covered. It was decided to limit the corpus to Bridge Team Communication covering standard navigational tasks only. These standard tasks comprise verbal exchanges on navigational situations such as manoeuvring operations in ports, fairways and the open sea, discussions on meteorological conditions as the influence of wind and currents and assessing prevalent traffic situations. The latter topic dwells on the assessment of ship traffic in the vicinity and on the compliance with the Collision Avoidance Regulations (International Maritime Organization 1972). Tasks on pilotage, berthing, cargo-related work and emergency situations have not been included into the text corpus as they might have compromised the overall homogeneity of the sampled speech events[27]. Verbal exchanges in standard navigational tasks constitute the bulk of bridge team communication and may be used at a later stage as a benchmark value for studying communicative behaviour in specific or exceptional situations as in the case of emergencies or contingency events.

Following the preliminary (pilot) studies, the speech data was collected by audio-

27 See chapter *4.3. Corpus representativeness*.

recording students of Nautical Sciences who had given their informed consent in compliance with the Social Sciences Human Research Ethics regulations of the University of Tasmania (Ethics reference number H0013035). Although university students, all participants had a considerable professional experience on board sea-going ships[28]. No participants withdrew their consent during or after the recordings thus eliminating any attrition effects.

In total, the developed text corpus consists of naturalistic verbal interaction by 40 students (i.e. 20 German and 20 Irish students) who were audio-recorded in 20 exercises which took place in the years 2013 and 2014. The non-intrusive recordings were transcribed manually by the researching student. In order to ensure a reliable transcription process, audio segments of 15 minutes each were loaded into the Audacity computer software which allows the insertion of text tracks. The transcribed texts were inserted into these text tracks below the respective communicative exchange. By doing so, subsequent checks and modifications of the transcribed text are considerably simplified as the transcribing student could listen to the audio files and read along the text simultaneously. Metadata were collected for all participants by means of questionnaires, including their the participants' age, gender, mother tongue, sea experience, self-rated English proficiency and work environment during the time they had served on board sea-going ships.

The transcript was validated by the student's supervisors who random-checked approximately 20% of the resulting text corpus in terms of words. Ambiguous words and expressions were discussed and corrected. Ambiguities which could not be resolved were marked as unintelligible words and disregarded in the linguistic analyses. The transcripts of the exercises were trimmed to a duration of 60 running minutes each to make them cover the standard navigation tasks only.

28 Due to international regulations, university studies of Nautical Sciences include a one-year practical experience as a cadet on sea-going ships.

4.2. Processing of primary data (corpus building)

Nakamura and Sinclair (1995) and McEnery, Xiao and Tono (2006, p. 14) highlight the importance of corpus building based on external criteria in order to ensure that the naturalistic linguistic samples remain independent of the corpus selection. This selection might suffer from a problematic circularity when selecting text samples with a predefined linguistic distribution (Atkins, Clear and Ostler 1992, p. 5-6; Biber 1993, p. 256).

By selecting the setting of the recorded exercises as described in chapter *4. Data collection and data processing* the Bridge Team Corpus has been developed based on external criteria. The overall aim was to provide the highest possible degree of similarity in the tasks to be carried out and to curb the recordings to an identical time frame. While an effort was made to reduce any possible influence on the research outcome by these criteria, other factors needed to be tested for a possible bias. For this purpose, participants were asked to provide demographic information including:

- their gender and age,
- their work experience on board sea-goings ship,
- the composition of these ship's crews, and
- the language predominantly spoken on board.

Upon conducting the pertinent analyses on these factors to determine a possible bias it was found that none of the differences in the participants' demographic data influenced the linguistic variables studied as dependent variables in a statistically significant manner[29].

29 For a detailed summary of the conducted analysis, see chapter *10.4 Paper IV*, section *6. Data analysis*.

4.3. Corpus representativeness

The specialised text corpus developed as an integral part of this doctoral research aims to be representative for the genre of Bridge Team Communication as produced by native and non-native speakers of English, in line with Leech's (2006, p. 13) argumentation that "[a] corpus is thought to be representative of the language variety it is supposed to represent if the findings based on its content can be generalized to the said language variety".

Biber (1993, p. 243) suggests a statistical definition of corpus representativeness which "refers to the extent to which a sample includes the full range of variability in a population" while McEnery and Wilson (2001, p. 105) emphasise that corpus building necessarily involves "dealing with a sample of a much larger population". Following Biber's suggestion, the population for which inferences are made in this research include future nautical officers from Ireland and Germany who are in their final year of training and who have been recorded in full-mission simulation exercises comprising standard navigational tasks. However, the speech samples collected from the Irish native speakers may be generalised to all native speakers of English, and to a certain extent the German non-native speakers may represent a population of all non-native speakers of English with a similar education and training background and professional on-board exposure[30].

These two claims may be arguable as native speakers of English vary considerably in their linguistic output across different countries and regions. However, given the specific speech events on a ship's bridge and the high degree of standardised maritime vocabulary native speakers are expected to avoid any regional idiosyncrasies to improve comprehension and thus achieve a better comprehension by other native speakers and non-native speakers of English. The same applies to non-native speakers which also need to limit their language proficiency to produce those utterances which aim to facilitate the exchange of professional information. The limited use of lexical structures in Bridge Team Communication in the speech samples soon leads to a saturation point beyond which hardly any new lexical items are introduced[31].

30 Chapters *10.3. Paper III* and *10.4 Paper IV* contain a more detailed discussion on the speech samples' possible generalisation to other speaker populations.

31 See chapter *4.3.2. Corpus closure* for further information on the saturation of lexical items.

In terms of the communicative settings covered, Biber's representativeness requirements have been meet in the developed speech corpus for standard navigational tasks including operations in ports, fairways and the open sea, discussions on meteorological conditions and given traffic situations (e.g. ship traffic in the vicinity and compliance with the Collision Avoidance Regulations). 4.3.1. Corpus size

The sample selection process for the Bridge Team Communication corpus is based on external criteria as stated in chapter *4.2. Processing of primary data (corpus building)*. For this reason, the recording time is the chosen criterion to determine the sample length. All of the 20 transcribed exercises have a duration of 60 minutes each. However, they result in quite different text lengths, with the native-speaker sub-corpus containing a total of 63,871 word tokens and the non-native speaker sub-corpus a significantly lower number of 43,019 word tokens. As the object of study is spontaneous verbal communication, it is quite normal to observe that some participants speak more than others. The number of produced utterances (and therefore word tokens) is thus studied as one of the dependent variables. While the text length of the different team transcripts was allowed to vary considerably, an effort was made to reduce the number of independent variables to a minimum, the most notable of which is the nativeness of the speakers[32].

The question arises whether the specialised text corpus containing 106,890 word tokens is sufficiently complete to be considered maximally representative for its population.

4.3.2. Corpus closure

A specialised text corpus can be considered representative if it contains a sufficiently high number of the linguistic features to be studied and these features appear "to be finite or […] subject to very limited variation beyond a certain point" (McEnery, Xiao and Tono 2006, p. 16). This particular point is termed *closure* and describes a stage at which a "particular feature in a variety of language is becoming finite" (Baker, Hardie and McEnery 2006, p. 33), i.e. the observed feature does not change significantly by increasing the sample size. As Baker, Harder and McEnery put it, "[t]he more a corpus approaches closure, the more it approaches being completely representative of a

32 For an in-depth analysis of the dependent variables examined and possible biasing effects of (independent) control variables on the results, see chapter *10.4 Paper IV*.

language (or language variety). So, the more representative a corpus becomes, the lower the likelihood that new words, phrases or grammatical rules will be found" (ibid).

In naturally occurring speech, a complete *lexical closure* is highly unlikely as it would mean that not even one new word type would be added at some point in the future. On the other hand, even if lexical closure is reached with a particular sample size, this does not necessarily indicate that by increasing the sample size further no new types would surface. Because of this uncertainty, McEnery and Wilson (2001, p. 176) suggest that "a clearer indication of closure might be given by plotting the growth of the [...] lexicons as graphs" and compare the evidenced tendency "towards premature closure" (ibid).

While this proposal is useful for comparing tendencies across different text corpora, it does not suggest a specific ratio as a trade-off value at which the degree of lexical closure might be considered sufficiently high to cater for a representative sample. For this reason, Baker Hardie and McEnery (2006, p. 33) suggest that lexical closure may be "defined as the point in a corpus beyond which the number of new lexical forms seen in every additional 1,000 tokens begins to level off at a rate lower than 10 per cent".

While *lexical closure* is limited to word types, *part-of-speech (or POS) closure* refers to the syntactic functions of words, which may vary according to their grammatical use. The word *anchor*, for example, can be used as a noun (*an anchor*) but also as a verb *to anchor*). For this reason, a sample corpus might reach a very high degree of lexical closure without even getting close to its part-of-speech saturation point.

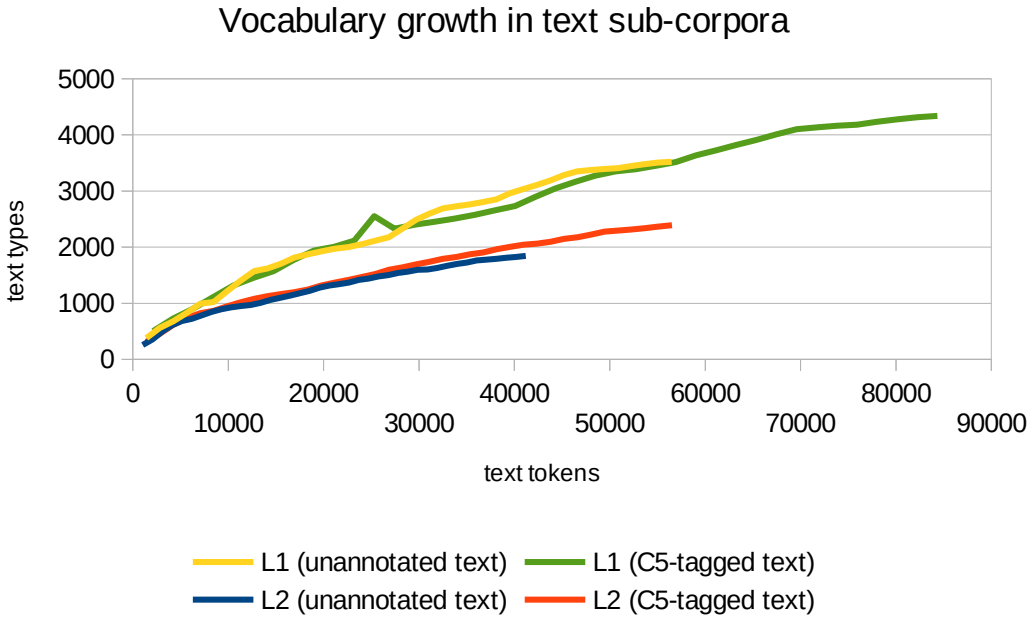

figure 4: Vocabulary growth in text sub-corpora (own illustration)

Following the methods proposed by McEnery and Wilson (2001), and Baker, Hardie and McEnery (2006), the Bridge Team Communication corpus has been tested for lexical and part-of-speech closure both graphically and by calculating the number of new types when successively adding 1,000 tokens. The corpus' tendency towards lexical and part-of-speech closure is displayed in figure 4, where L1 describes the figures for the native speaker sub-corpus and L2 for the non-native speaker sub-corpus. Figure 4 does not only depict a relatively gentle increase of new word types over the total text length, it also evidences that the lexicon growth starts levelling off towards the end of the curves[33].

In order to corroborate the findings of the graphical lexical and POS growth curves, the number of new types relative to each increase of 1,000 tokens has been calculated. Table 6 displays the values for the last 1,000 token increase. The values clearly demonstrate lexical and POS closure for both the L1 and L2 sub-corpora separately and together as one combined corpus (L1 + L2).

Given the very high degree of closure presented by the Bridge Team Corpus for lexicon and POS growth figures, the corpus is considered to be sufficiently representative in

33 For a comparison and analyses of the total Bridge Team Corpus with reference corpora, see chapter *10.3. Paper III,* section *3 Data analysis*.

terms of the linguistic features studied.

(sub-) corpus	**closure type**	**new types per 1000 tokens**	**degree of closure** (acc. to Baker et al)
L1	lexical	12	1.2%
L1	part-of-speech	12	1.2%
L2	lexical	17	1.7%
L2	part-of-speech	18	1.8%
L1+L2	lexical	13	1.3%
L1+L2	part-of-speech	10	1.0%

table 6: Analysis of lexical and part-of-speech closure

4.3.3. Dispersion

Another method to estimate the degree of representativeness is to measure the dispersion of the studied linguistic features in a text corpus. According to McEnery and Wilson (2001, p. 80) "[d]ispersion is a measure of how evenly distributed the occurrence of a feature is in a text or corpus – for example, whether its appearance is restricted mainly to a few places or whether it occurs much more widely". Baker, Hardie and McEnery (2006, p. 59-60) describe dispersion measures as a means to "determine whether a term is equally spread throughout a text or occurs as a central theme in one or more parts of the text". It is assumed that the highest degree of representativeness is achieved by a specialised text corpus if it presents homogeneous dispersion figures for the text variety or genre under scrutiny. In statistical terms, a homogeneous dispersion indicates that the sampled texts are indeed representative of the population for which inferences are made.

Baayen studies dispersion figures as a possible cause for overestimation bias and reasons that

> "[a] word is underdispersed if its observed dispersion is smaller than its expected dispersion. Since underdispersed words occur in fewer text chunks than expected under chance conditions, they are the words the tokens of which occur concentrated in particular parts of the text, instead of being spread out evenly throughout the text. Hence, the underdispersed words are the words that, if our hypothesis is correct, should be responsible for the overestimation bias."
>
> (Baayen 2001, p. 164)

In order to compute lexical dispersion characteristics, Baayen suggests a dispersion analysis calculation based on Monte Carlo simulations which compare the observed dispersion of words with their probability under chance conditions (ibid).

Following Baayen's methodology, a number of Monte Carlo-based analyses has been performed on the developed Bridge Team Communication corpus. In total, 5,000 permutation runs were performed on 100 text chunks for all words included in the text corpus as a whole and separately for the native speaker (L1) and the non-native speaker (L2) sub-corpora. The analyses were undertaken on the transcribed, unannotated text as well as on the processed text including its C5 part-of-speech tagging. Subsequently, the distributions of the observed and the expected dispersion measures were analysed by means of the non-parametric Mann-Whitney U test and the respective Probability of Superiority (PS) effect sizes were computed. Table 7 summarises the findings of the dispersion analysis.

The statistical analyses carried out on the unannotated and on the POS-tagged corpus and its two sub-corpora all produce a PS effect size which is extremely close to the value of 0.50 which would indicate a random distribution of the observed effects. Both the original and the processed text present dispersion figures which indicate a highly homogeneous lexical spread, thus indicating a very representative sample of its population with a minimal overestimation bias caused by underdispersion.

(sub-) corpus	**version**	***p* value**	**PS**
L1	unannotated	0.002	0.48
L1	C5-tagged	0.006	0.48
L2	unannotated	0.002	0.47
L2	C5-tagged	< 0.000	0.48
L1+L2	unannotated	< 0.000	0.47
L1+L2	C5-tagged	< 0.000	0.48

table 7: Analysis of lexical dispersion

4.4. Corpus mark-up

A text corpus may only fulfil the desired "particular representative function" (Leech 1991, p.11) outlined in chapter *4. Data collection and data processing* if the raw text is complemented with additional, "extra-textual information" (McEnery and Wilson 2001, p. 39) in a comprehensive and systematic manner. McEnery, Xiao and Tono (2006, p. 22) describe this *corpus mark-up* process as "a system of standard codes inserted into a document stored in electronic form to provide information about the text itself and govern formatting, printing or other processing". The aim of corpus mark-up is to deliver "relatively objectively verifiable information regarding the components of a corpus and the textual structure of each text" (ibid, p. 29). Baker (2006, p. 38) suggests that "corpus builders employ some form of annotation scheme to their text files, however brief, in order to aid analysis and keep track of the structure of the corpus".

The integration of extra-textual information is necessary to contextualise the unprocessed text, or as McEnery, Xiao and Tono (2006, p. 22) put it, "to relate the specimen to its original habitat". The authors stress the fact that "contextual information is important in recovering the situation in which a particular corpus sample was produced" (ibid, p. 74). This statement corroborates earlier findings by Crowdy (1993, p. 264) who states that "[d]etailed discourse analysis cannot be satisfactorily carried out without some knowledge of the situation or context in which the discourse has taken place".

For the purpose of disambiguation and contextualisation a questionnaire was developed during the design phase of the Bridge Team Communication corpus to collect the extra-textual information required for studying the identified research variables[9]. In order to identify the different recordings, the corpus contains information on the place and date of the exercises, and identifiers make reference to the original audio files and to the role the participants were assigned in the exercise (i.e. shipmaster or first officer). Gathered meta-data include the participants' age, gender, mother tongue, sea experience, self-rated

9 The questionnaire is included in chapter *11.3 Appendix 3: Survey of participants' demographic data*.

English proficiency and work environment during the time they had served on board sea-going ships[10].

The example displayed in table 8 outlines the corpus mark-up structure based on the questionnaire.

4.5. Corpus annotation and encoding

While the term *mark-up* refers to extra-textual information added to the text corpus, the terms *annotation* and *encoding* describe the inclusion of textual analyses of the original, unprocessed text. Leech (1997, p. 2) defines corpus annotation as the act of "adding [...] interpretative, linguistic information to an electronic corpus of spoken and/or written language data" while Baker, Hardie and McEnery (2006, p. 67) see encoding as "an analysis of some feature at the discourse, semantic, grammatical, lexical, morphological or phonetic level".

To provide for a quantitative analysis of the Bridge Team Communication, data fields for processed speech data have been added. These fields include part-of-speech and SMCP key word tagging, calculations of linguistic ratio and frequency variables at an utterance level as well as pragmatic markers.

The example displayed in table 9 contains the encoded speech data for the utterance listed in table 8.

Here, the fields <*time in sec[onds]*>, <*time difference*> and <*communication segment*> refer to the time difference between the individual utterances as analysed in paper I[11].

10 For a detailed overview of the sampled participants and a more detailed explanation of the meta-data, see chapter *10.4 Paper IV,* section *6. Data analysis*.

11 See chapter *10.1. Paper I,* section *3.1 Definition of independent communication segments*.

< utterance ID>:	<145>
<speaker>:	<shipmaster>
<team size>:	<2>
<audio file>:	<2013-03-21 Elsfleth Weser.wav>
<place>:	<Elsfleth>
<vessel>:	<Weser>
<age>:	<27>
<gender>:	<male>
<experience>:	<13>
<position>:	<cadet>
<last work>:	<2010>
<start English>:	<1991>
<self-rating>:	<good>
<nationality>:	<German>
<mother tongue>:	<German>
<crew>:	<always>
<officers>:	<English>
<raw text>:	<if the current is going this way we have to cross here, to get, to get there>

table 8: Mark-up structure of Bridge Team Communication text corpus

<time in sec>:	668
<time difference>:	1
<communication segment>:	1
<POS-tagged>:	if_CJS the_AT0 current_NN1 is_VBZ going_VVG this_DT0 way_NN1 we_PNP have_VHB to_TO0 cross_VVI here_AV0 ,_PUN to_TO0 get_VVI ,_PUN to_TO0 get_VVI there_AV0
<SMCP-tagged>:	if the SMCP_current is SMCP_going this SMCP_way we have to SMCP_cross SMCP_here, to SMCP_get, to SMCP_get SMCP_there
<word count>:	17
<content words>:	8
<lexical density>:	0.47
<key words>:	8
<key word density>:	0.47
<GDI>:	0.85
<spdi>:	0.93
<directive:>	0
<commissive:>	1
<ambiguous utterance:>	0
<reception; no risk:>	0
<reception; residual risk:>	0
<reception; high risk:>	0

table 9: Annotation structure of Bridge Team Communication text corpus

4.5.1. Part-of-speech tagging

In order to identify the different grammatical word classes uttered by the recorded speakers, the data field <*POS-tagged*> contains the original utterance provided with annotated semantic tags. The tagging operation itself was carried out with the CLAWS4 part-of-speech tagger which was also used to tag the British National Corpus (Garside 1987; Leech, Garside and Bryant 1994; Garside 1996; Garside and Smith 1997). For tagging the Bridge Team Communication corpus, the C5 tag set was used[37]. Although the CLAWS4 tagger is reported to achieve a consistent accuracy of 96 to 97 percent (Garside 1996; Leech, Garside & Bryant 1994) this result was further improved by checking and correcting the Bridge Team Corpus tags manually.

Other methods to discriminate grammar structures such as syntactic parsing were considered during the preliminary (pilot) studies carried out prior to the collection of the Bridge Team Corpus. Although syntactic parsing is a valuable method to extract grammatical structures of utterances, it was found that given the relatively simple grammar structures of the studied verbal exchange, no significant differences could be established between native and non-native speakers of English (Mann-Whitney U, p=0.001). For this reason, the methodology of syntactic parsing was has been disregarded in this research.

By counting different word classes in two different groups of speakers (namely native and non-native participants) numerical sample distributions are created which can be analysed statistically. *Lexical density*, for example, is a ratio obtained by dividing all content words (nouns, full verbs, adjectives and adverbs) by the total word count. The data fields <*word count*>, <*content words*> and <*lexical density*> are used for this purpose[38].

4.5.2. Key word tagging

While the used part-of-speech tagging was carried out by means of a defined tag set and existing computer software, no tool was available for tagging SMCP key words. For this reason, all content words were extracted from the Standard Marine Communication Phrases by means of Quantitative Content Analysis (Boettger and Palmer 2010; Frey,

37 See chapter *10.2. Paper II*, section *3.1 Data Sampling and Definition of Grammar Diversity*.

38 As outlined in chapter *10.1. Paper I*, section *3.3. Weighted analysis*.

Botan and Kreps 2000, p. 236-244; Riff, Lacy and Fico 2014; Rourke and Anderson 2004), and a computer programme was written by the researching student to annotate these SMCP key words and calculate the key word density accordingly[39]. For a better recognition rate and thus a higher reliability, differing word forms were studied and added manually (e.g. cargo – cargoes, carry – carried – carrying) until all content words included in the SMCP were reliably tagged by the software.

By dividing the number of SMCP-tagged words calculated in the <*SMCP-tagged*> field by the total word count included in the <*word count*> field, the resulting key word density ratio is inserted into the corresponding <*key word density*> data field.

The lexical and SMCP key word densities are analysed as dependent variables on an utterance level in papers I, III and IV[40].

4.5.3. Tagging of grammar diversity markers

The data fields <*GDI*> (i.e. Grammar Diversity Index) and <*spdi*> (i.e. special POS diversity index) are used to compute an observed part-of-speech diversity in each utterance against an expected value[41]. By comparing observed and expected *spdi* values, inferences can be made on the complexity of the grammar classes used by the different speakers.

The grammar diversity index and the special part-of-speech diversity index are defined in paper II, and analyses on values computed as dependent variables for the Bridge Team Communication are performed in papers III and IV.

4.5.4. Tagging of pragmatic markers

While encoding utterances for the data fields mentioned above was carried out by computer software with subsequent manual checks and improvements by the researching student, the tagging of pragmatic markers follows a Corpus Pragmatics approach. Here, frequencies of linguistic features are identified vertically (i.e. by means of Corpus

39 See chapter *10.1. Paper I*, section *3.3. Weighted analysis*.

40 See chapters *10.1. Paper I*, *10.3. Paper III* and *10.4 Paper IV*.

41 For a detailed description on POS diversity see chapter *10.2. Paper II,* section *3.1 Data Sampling and Definition of Grammar Diversity.*

Linguistic methods) in order to be analysed horizontally by employing a Pragmatics approach[42].

The pragmatic markers <*directive:*> and <*commissive:*> are used to identify these two illocutionary speech acts based on a combined search for speech act clues while the data fields <*ambiguous utterance:*>, <*reception; no risk:*>, <*reception; residual risk:*>, <*reception; high risk:*> are used to identify three discrete risk categories in the respective perlocutionary acts[43].

5. Data analysis

After completing the annotation and encoding of the Bridge Team Communication corpus the identified and computed variables have been analysed by means of statistical methods. These analyses follow the classification of quantitative research provided by Dörnyei (2007)[44].

- *Focus on numbers*: Linguistic observations have been translated to numerical values by counting occurrences on an utterance level. This method leads to statistical distributions for individuals, teams, groups (e.g. L1 and L2 speakers) or communication (time) segments.

- *A priori categorisation*: The applied measurement instruments are based on part-of-speech tagging, with the exception of the horizontal, pragmatic analysis conducted in paper V. The C5 and SMCP tag sets[45] were used in all studies delivering a consistent and reliable way of quantifying linguistic observations.

- *Variables instead of cases*: The analyses were carried out on independent and dependent variables defined *a priori*. The independent variables included all meta-data collected through the participant questionnaires[46], namely their

42 For a detailed description of the Corpus Pragmatics concept and an overview of relevant literature, see chapter *10.5 Paper V*, section *2 Methodology*.

43 For a detailed description of the employed speech act categories, see chapter *10.5 Paper V*, section *3 Identification of speech acts*.

44 See chapter *3.1. Methodological framework*.

45 For a detailed explanation of the used C5 tag set, see *4.5.1. Part-of-speech tagging*.

46 The questionnaire is included in *11.3 Appendix 3: Survey of participants' demographic data*.

biographic data (i.e. gender, age, nationality, mother tongue) and information concerning their exposure to real-life working conditions on board sea-going ships (i.e. time at sea, position on board, last work experience, working language on board, rating of own English skills). The studied dependent variables include the data fields listed in *table 9: Annotation structure of Bridge Team Communication text corpus*, i.e. communication segments, word counts, lexical density, key word density and special part-of-speech diversity as well as speech act and risk analysis markers.

- *Statistical analysis*: The numerical distributions provided by the outlined approach are analysed by statistical methods to identify potentially significant differences between two samples based on the independent variables listed above. Upon detecting a statistically significant difference, the magnitude of the effect is calculated by means of appropriate effect sizes. These effect sizes deliver a numerical value which can be compared to other observations and ranked accordingly.

- *Standardised procedures to assess objective reality*: The measurement instruments were selected to provide the highest possible degree of objectivity[47]. Quantitative instruments have also proven to be highly reliable thus reducing observer bias considerably[48].

- *Inferences are made on samples*: Generalisations are made on the basis of a sampled group of individuals. By creating statistical distributions on groups, the effect of individual variations on the overall result is reduced due to a regression to the mean effect.

47 See chapter *3.6. Epistemology*.

48 These aspects are further discussed in chapter *6. Measures of research quality and robustness*.

5.1. Frequencies and proportions

In this research, a quantitative profile of Bridge Team Communication is modelled by means of frequentist inferencing whereby inductions are made on frequencies or proportions computed for studied samples (Mayo and Cox 2006; Vasishth and Nicenboim 2016; Wagenmakers, Lee, Lodewyckx and Iverson 2008). This approach is in accordance with Koplenig (2017, p. 2), who asseverates that "[t]he main idea behind statistical frequentist inference is to use the distributional information from a sample of objects to estimate the characteristics of the unknown population from where the sample was taken". According to Bod (2003, p. 14), "under the frequentist interpretation, the probability of an event is interpreted as its relative frequency in a series of experiments". Given that the studied sample is representative of its target population, frequentist inferencing provides the means to establish a quantitative linguistic model which can be used for comparing observed communication structures against predicted model values.

Statistical analyses in this doctoral research are carried out on raw frequencies, e.g. the proportion of nautical key words to the total word count per utterance[49]. Given the sample size is sufficiently representative[50], the distribution of the computed frequencies can be considered representative for the linguistic genre of Bridge Team Communication as the "unknown population from where the sample was taken" (Koplenig 2017, p. 2).

5.2. Statistical methods

Frequentist statistics adheres to a predefined inference framework which includes statistical hypothesis testing. This research has defined a number of null hypotheses (H_0) with their respective alternative hypotheses (H_a), and appropriate statistical methods have been employed to accept or reject the null hypotheses. The employed methods include parametric (e.g. analysis of variance or ANOVA testing) and non-parametric hypothesis testing (e.g. Kruskal-Wallis and Mann-Whitney U tests), depending on the distribution patterns of the studied variables (standard Gaussian or non-standard

49 For more information on calculating proportions for each utterance, see chapter *10.1. Paper I*, section *3.3. Weighted analysis*.

50 For further details on the sample size, see chapter *4.3. Corpus representativeness*.

distribution). This latter aspect was checked with tests for normality (e.g. Shapiro-Wilk test) and homoscedasticity (e.g. Levene's test for homogeneity of variances).

For statistically significant findings in hypothesis testing on metric variables, the magnitude of the effects has been quantified with the Probability of Superiority (PS) effect size described by Grissom and Kim (2005, p. 98) as "the probability that a randomly sampled member of population *a* will have a score (Y*a*) that is higher than the score (Y*b*) attained by a randomly sampled member of population b." The PS effect size neither assumes a normal nor a homoscedastic distribution of the studied variables and is therefore very useful for dependent linguistic variables which often do not present a Gaussian distribution[51].

In the case of categorical variables, the magnitude of the effect has been calculated by means of the phi coefficient (Φ) as recommended for naturalistic research on dichotomous variables by Grissom and Kim (2005, p. 249)[52].

Apart from hypothesis testing this research has also employed other statistical methods such as regression analyses to model the relationship between two dependent metric variables (e.g. the expected number of different grammar classes per utterance length). Further details on the applied statistical methods are included in the respective research papers.

6. Measures of research quality and robustness

The adopted quantitative methodology aims to achieve a means of analysing authentic spoken communication without the need for an introspective judgement. Even if McEnery and Wilson (1996, p. 14) hold that "[c]orpus based observations are intrinsically more verifiable than introspectively based judgement" the conducted research still needs to prove its reliability and validity in a transparent manner.

51 For more information on the Probability of Superiority (PS) effect size, see chapter *10.3. Paper III*.

52 For more information on the phi coefficient (Φ), see chapter *10.5 Paper V*.

6.1. Reliability

According to Jones (2012, p. 351) "reliability equals consistency". This brief statement refers to the very notion of *reliability* in academic research which aims to answer the question asked by Brown (2013, p. 492): "To what degree would the results be the same if the instrument were administered repeatedly?"

McEnery and Wilson (2001, p. 14) argue that "[c]orpus-based observations are intrinsically more verifiable than introspectively based judgements". The high degree of reliability is achieved by the observations' *a priori* categorisation stated by Dörnyei (2007)[53]. Nevertheless, even quantitative Corpus Linguistics methods cannot ensure that the reported results are totally (i.e. 100 percent) reliable. In order to assess the reliability of an adopted research approach, Brown (2013, p. 492) thus encourages researchers to "demonstrate (1) the reliability of the instruments used in their studies and (2) the reliability of the results of their studies".

6.1.1. Reliability of instruments

Research instruments are reliable if they provide consistent results. The adopted methodology should thus optimally deliver identical results when applied to the same data again.

In this research, a small degree of uncertainty remains as to the transcribed audio samples. Transcriptions were carried out manually which bears the possibility of misunderstanding or inaccuracies. To minimise these risks, a rigorous effort was made to ensure the correctness of the transcribed verbal interaction. To achieve a high level of reliability, the researching student double-checked all transcripts by listening to them again several times after the transcripts were finished. Remaining mistakes were corrected at this stage. The transcripts were subsequently rechecked by the student's supervisors[54]. Remaining unintelligible words that could not be ascertained due to overlapping speech or other interferences were marked as *unintelligible* so that an analysis of their impact on the overall result could be carried out (see next chapter).

53 See chapter *5. Data analysis* for more information.

54 See chapter *4.1. Collection of primary data (sampling)* for more information on the transcribed data.

Hence, the outlined procedure of transcribing the recorded verbal interaction by bridge team members achieves a sufficiently high reliability as to provide consistent results.

Upon completion of the transcription, the raw data was encoded by means of part-of-speech (POS) and key word tagging. The used tagging software for identifying word classes provides a reported reliability of 96 to 97 percent[55]. In order to improve this result further, the computer-tagged utterances were checked by the researching student, and mistakes were corrected manually with the aim to achieve the highest possible accuracy. Unintelligible words received a specific POS marker so that they could be identified and checked for their distributional patterns later.

Another source for reducing reliability are statistical errors which occur as an integral part of hypothesis testing. A null hypothesis might be true but is erroneously rejected, hence a statistically significant difference is assumed where no such difference exists. In order to estimate the probability of this so-called *type I error* the significance level has been stated in all statistical testing. For the dependent linguistic variables under scrutiny, a value of $p < 0.00$ is assumed to provide a sufficiently reliable result as a *type I error* can only be expected in one analysis out of 1,000 (Brown 1988; Hatch and Lazaraton, 1991). By calculating the respective effect size for each dependent linguistic variable the magnitude of the observed effects are quantified, and inferences can be made with regards to a possible impact on the research results. Calculating the magnitude of the observed effects has also provided a means of cross-checking the values resulting from hypothesis testing.

6.1.2. Reliability of results

The adopted research paradigm of frequentist inference[56] assumes that a given sample provides valid data for estimating values of the population the specific sample is taken from. As the population is unknown, any sampling process is subject to *sampling error*, i.e. a difference between a statistic observed in the given sample and the population these values are projected on. In order to estimate a possible sampling error in this research, statistical modelling has been undertaken on the corpus representativeness, its size and the possible saturation of dependent linguistic variables (i.e. corpus closure)[57].

55 See chapter *4.5.1. Part-of-speech tagging* for more information.

56 For more details on frequentist inference, see chapter *5.1. Frequencies and proportions*.

57 For more details on this modelling, see chapter *4.3. Corpus representativeness*.

In order to estimate the possible impact of unintelligible words on the SMCP key word ratio a probabilistic approach was adopted whereby the wildcard characters identifying unintelligible words were assumed to be key words based on their most likely distribution. For this reason, a linear regression analysis was undertaken with the aim to determine the number of SMCP key words for a given number of words per utterance, resulting in $f(x)=0.24x$ and a corresponding coefficient of determination of $R^2=0.82$.

For each utterance containing unintelligible words the probable number of key words was calculated so that a new key word distribution was created. The original and the probable key word counts were introduced into a contingency table and analysed by means of a chi-squared test of independence. X^2 resulted in p=0.06 so that the null hypothesis of no statistically significant difference between both key word counts is accepted.

As a cross check both key word distributions were also compared by means of a Mann-Whitney U test (MWU) which resulted in p=0.004. The corresponding Priority of Superiority effect size was PS=0.49. This value is extremely close to 0.50 which would indicate a random distribution in two samples.

The procedure was repeated in the same fashion for the content word distribution, resulting in $f(x)=0.45x$, $R^2=0.89$, $X^2<0.00$, $p_{MWU}<0.00$ and PS=0.48. In this case, the computed X^2 and p_{MWU} values both indicate a statistically significant difference between the two distributions ($\alpha=0.05$). However, the very low PS effect size of 0.48 is assumed to warrant a sufficiently similar distribution to minimise the bias introduced by unintelligible words.

The statistical calculations for determining vocabulary growth figures, type-token ratios and the special part-of-speech index are all based on content word and SMCP key word distributions. Given the extremely low magnitude of the effect sizes calculated for these distributions the impact of unintelligible words in the transcript can safely be disregarded.

Based on the outcome of the statistical modelling on sampling error and on content and key word distributions it is assumed that the achieved research results provide a sufficiently high “degree to which the results would be likely to reappear if the study were replicated under the same conditions” (Brown 2013, p. 492). At the same time, the

frequentist inferences made on the analysed sampled are assumed to deliver a high degree of representativeness of their target population, i.e. Bridge Team Communication as a linguistic sub-genre of Maritime English.

6.2. Validity

While the term reliability refers to the consistency of the employed research instruments and their results, the concept of validity describes "the extent to which [an instrument] measures what it is supposed to measure *and nothing else*" (Heaton 1975, p. 153). Hammersley (1987, p. 69) provides a similar definition for validity: "An account is valid or true if it represents accurately those features of the phenomena, that it is intended to describe, explain or theorise". Validity has been referred to as the "degree of approximation of 'reality'" (Johnston and Pennypacker 1980, pp. 190-191).

Validity may be considered one of the most important aspects in language *proficiency* testing where is has been discussed extensively (Bachman and Palmer 1996; Fulcher and Davidson 2012; McNamara 2000; Weir 2005). In an assessment of naturalistic language *performance,* the observations carried out need to represent the specific linguistic genre and speech event under scrutiny both internally and externally in line with Brown (2013, p. 493):

> "Internal validity is the degree to which the results of a study can be accurately interpreted as meaning what they appear to mean. External validity is the degree to which the results of a study are contrived or artificial, or put another way, the degree to which the results apply to the outside world."
>
> (Brown 2013, p. 493)

6.2.1. Internal validity

Kerlinger (1964, p. 430) refers to the concept of validity by asking if we "are measuring what we think we are". In order to achieve a high degree of internal validity in the research results, the developed Bridge Team Communication corpus is analysed by means of ratios computed on counted word category frequencies. These categories include the constructs of lexical and key word density, vocabulary growth, part-of-speech diversity, cognitive words and the speech acts of commissives and directives. While the lexical and grammatical frequency count hardly includes any abstraction, the

constructs of lexical tokens expressing cognitive load levels and speech acts require a more detailed explanation. Both constructs are latent variables which cannot be measured directly. For this reason a careful selection of proxy variables is paramount to measure what they are "supposed to measure" (Heaton 1975, p. 153). This research into Bridge Team Communication builds on existing basic research which has validated the use of linguistic variables for measuring cognitive load[58] to achieve a sufficiently high validity to make inferences on this construct for the observed participants in bridge team simulation exercises. By adopting a methodology which has been tested and validated successfully by previous research, an Applied Linguistics approach is followed by systematically applying existing knowledge "through activities with a group of individuals and observing how they work to, for example, enhance learning or improve a process" (Phakiti and Paltridge 2015, p. 11).

For studying the construct of speech acts, this research has also built on the findings of previous research[59]. Commissives and directives uttered in the observed exercises have been identified and analysed by following an existing approach, thus applying validated basic research to the speech event of Bridge Team Communication in an Applied Linguistics fashion.

6.2.2. External validity

While internal validity relates to the instruments and the process of measurement, the concept of external validity refers to the applicability of the results to a wider research area. Creating a linguistic profile of Bridge Team Communication through a series of quantitative variables is only valid is these variables indeed provide a representative picture of the studied speech event. From a quantitative perspective, the validity of the analysed variables is warranted by providing statistically significant results across different linguistic settings and sociolinguistic groups[60]. The adopted approach has proven to discern Bridge Team Communication from other speech events inside and outside a maritime setting. It also offers a high probability in classifying the register of

58 A detailed review of existing literature on measuring cognitive load levels is included in chapter *10.4 Paper IV*, sections *2. Definition and measurement of cognitive load* and *3. Measuring cognitive load by linguistic variables*.

59 A detailed review of existing literature on the Speech Act theory by Austin and Searle is included in chapter *10.5 Paper V*, section *2. Methodology*.

60 As demonstrated in chapters *10.3. Paper III* and *10.4 Paper IV*.

an analysed text as spontaneous spoken communication. Native and non-native speakers of English have been found to provide statistically significant idiosyncratic speech patterns.

The linguistic profile created by the selected variables delivers a combination of quantified observations which may also be used in other linguistic settings. All variables are independent from the text genre with the exception of the studied key words which are taken from the Standard Marine Communication Phrases (SMCP). An adaptation of the profiling methodology to other genres or linguistic settings would require a modification of the used key words. The pursued quantitative profiling provides a good replicability and transferability to related research areas, thus providing a good generalisability.

7. Published results

This chapter summarises the main findings of the five research papers which form the core of this doctoral research. Apart from presenting the key results of each paper, the chapter also presents their inter-relationship and the methodologies which have been applied.

The five papers can be classified into two separate groups, the first of which is formed by papers I and II which provide the conceptual framework of this research. The publications of the second group apply the methods outlined in group one to study Bridge Team Communication from different angles.

Papers I and II introduce and validate the adopted concepts and methodology. The first paper identifies a gap in existing research into maritime communication and introduces a method for analysing Bridge Team Communication lexically. The second paper defines a method to identify grammatical diversity in language production. In both papers the introduced methods are validated on small corpora including transcripts of verbal communication taken from the public domain.

In papers III, IV and V the methods introduced by the two conceptual papers are applied to study a Bridge Team Communication corpus of authentic verbal interaction by native and non-native speakers of English. The three applied research papers apply the methods

introduced in the first two publications to investigate the specific discourse community of bridge team members from different perspectives. These perspectives include sociolinguistic, psycholinguistic and corpus-pragmatic aspects. Together, they establish a linguistic profile which can be deemed idiosyncratic for the sub-genre of Bridge Team Communication.

Figure 5 displays the inter-relationship of the research papers.

7.1. Results of paper I

The first paper titled "*Information density in bridge team communication and miscommunication – a quantitative approach to evaluate maritime communication*" identifies a lack of "empirical tool[s] […] to estimate the information flow of bridge team communication in a systematic manner" (John, Brooks, Wand and Schriever 2013, p. 233). While researchers involved in the maritime domain agree on the importance of effective Bridge Team Communication for safe navigation, no quantitative method exists to evaluate naturalistic language performance in this specific domain. In order to fill this research gap, the paper examines lexical density and key word frequencies and introduces a communication index which provides for a lexical analysis of the exchanged information. The communication index is also apt for making inferences on the utterances' maritime idiosyncrasy.

The index is validated on the verbatim transcript of a navigational accident available in the public domain. The validation process includes a mixed-methods analysis in which the identified text segments are scrutinised by means of Discourse Analysis. Here, the resulting index values are compared qualitatively with the prevalent situation on board and decisions made by bridge team members. An in-group comparison of the verbal exchange between the ship's captain and her pilot results in very significant differences between the individual text segments. By relating the speech content of the different segments to the respective index values the presented quantitative approach to evaluate the effectiveness of the information exchange is validated.

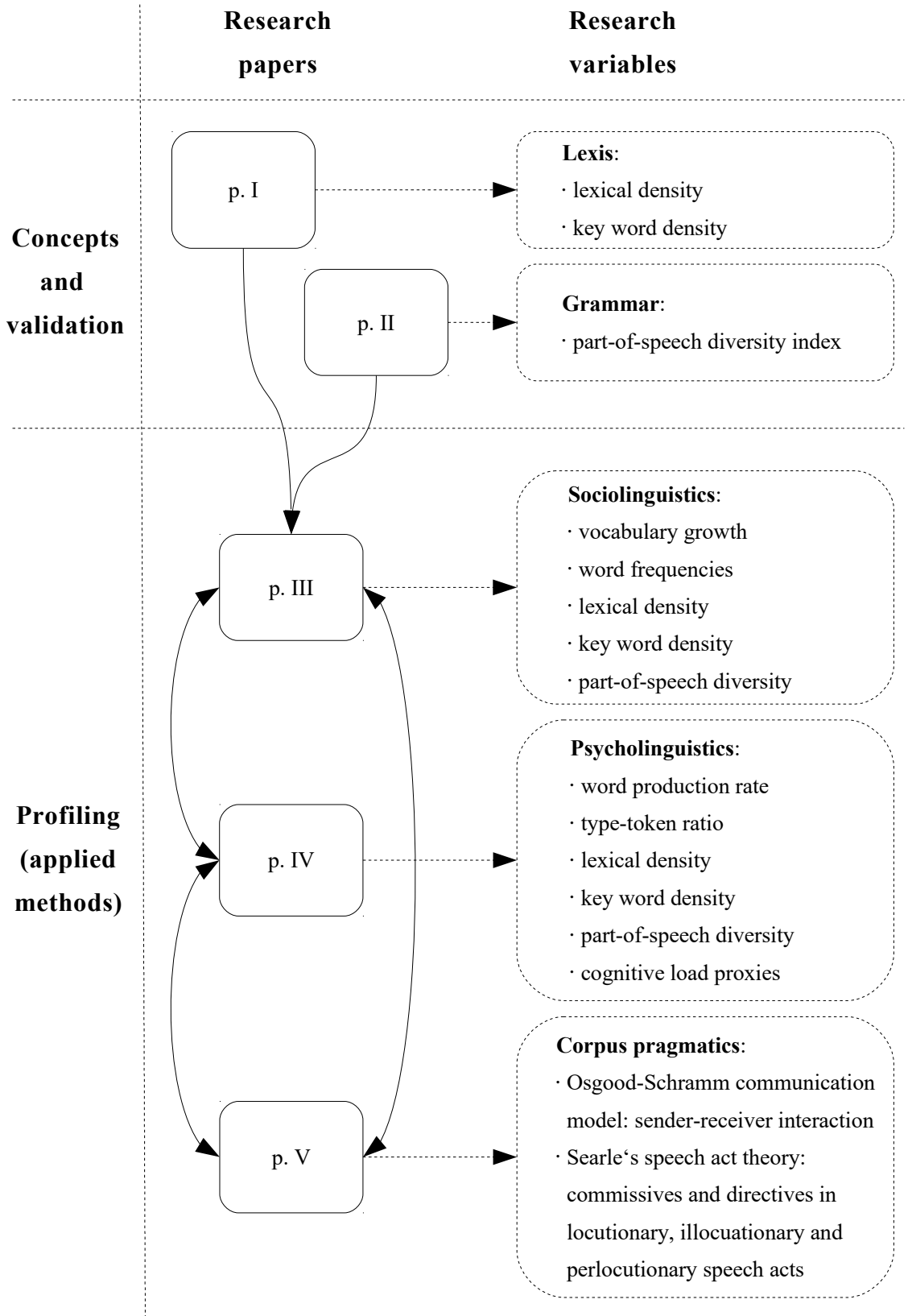

figure 5: Inter-relationship of research papers (own illustration)

The introduced methodology demonstrates that the "quality of the information content in bridge team communication is measurable" (John, Brooks, Wand and Schriever 2013, p. 242) by means of a "quantitative methodology to calculate and weight utterances for evaluating its information content" (ibid, p. 242). The intensity of the information exchanged amongst bridge team members varies over time (see figure 6) and reaches its highest values when unambiguous questions are answered (segment 13), a navigational situation is assessed (segment 27) and rudder commands are given and answered in a closed-loop communication (segment 29).

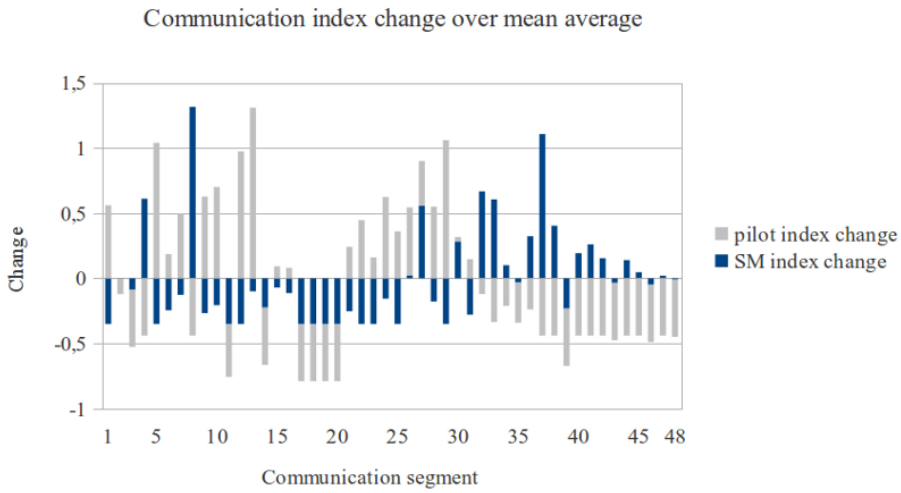

figure 6: Communication index change over mean average (John, Brooks, Wand and Schriever 2013, p. 240)

This paper sets the methodological framework for all subsequent publications. It introduces the concept of quantifying the studied linguistic variables by calculating frequencies or ratios at utterance level. These frequencies are computed by counting specific lexical word tokens or part-of-speech (POS) tags and dividing them by the respective utterance's total word count. The ratio calculation leads to a distribution which is subsequently analysed by means of appropriate statistical analyses. Hence, a frequentist approach is adopted for making inferences on probabilities based on collected speech samples.

7.2. Results of paper II

While paper I defines the methodological framework for analysing Bridge Team Communication lexically, paper II titled "*Lingua Franca and its Grammar Footprint: Introducing an Index for Quantifying Grammatical Diversity in Written and Spoken Language*" sets out to define an index "to quantify grammar diversity based on part-of-speech tagging (POS)" (John and Brooks 2014, p. 22). It therefore complements the lexical analysis introduced in paper I with an efficient tool to analyse the grammar structure of transcribed verbal interaction.

This methodological paper follows the frequentist approach adopted in paper I by re-using the same part-of-speech tags to calculate ratios at utterance level and determine a POS diversity index. The presented special POS diversity index delivers "the relative deviation of the grammar diversity value in an observed utterance from the expected number given by the PDI [POS Diversity Index] for a specific utterance length" (John and Brooks 2014, p. 33). It thus provides a robust methodology for comparing grammatical diversity for any given set of utterances, e.g. for a particular speaker or for a sociolinguistic group.

The presented index is validated on a small text corpus built by the researching student on the basis of freely available transcripts of radio interviews. The statistical analysis carried out on this corpus involves a between-groups comparison between the individual speakers and the radio presenter by means of an analysis of variance (ANOVA); the index itself is computed by carrying out a regression analysis. Following the methodology outlined in paper I, Discourse Analysis is employed to validate the calculated significant results in a qualitative manner.

Paper II demonstrates that

> "a between groups comparison can lead to an idiosyncratic spdi value for a given speaker. This grammar footprint can be compared with the values obtained for other speakers. It can also be used horizontally in research comparing different text types or discriminating spoken from written language, amongst other possibilities. […] By quantifying grammar diversity on the basis of part-of-speech (POS) tagging a simple and robust method has been presented which can be used along with other procedures using

> the same POS data for quantifying other linguistic features of written and spoken language"
>
> (John and Brooks 2014, p. 33-34)

Together, papers I and II provide the methodological framework for analysing Bridge Team Communication lexically and grammatically. Both the lexical and the grammatical analyses are based on ratios computed by dividing counts of word tokens or part-of-speech tags by the total word count of a particular utterance. Across all utterances, these ratios lead to statistical distributions which can be used for "unambiguous quantitative measures for comparative text analyses" (John and Brooks 2014, p. 34).

7.3. Results of paper III

Paper III titled "*Profiling maritime communication by non-native speakers: A quantitative comparison between the baseline and standard marine communication phraseology*" is the first paper of the applied research contributions which implement the conceptual framework and methods defined in papers I and II. All analyses performed in the second group of papers are undertaken on the Bridge Team Communication corpus described in chapter *4.2. Processing of primary data (corpus building)*.

This paper performs a genre analysis by comparing Bridge Team Communication by non-native speakers of English with two reference corpora outside the maritime domain and with the Standard Marine Communication Phrases (SMCP) text collection in order to identify "inherent structural patterns of nautical team communication" (John, Brooks and Schriever 2017, p. 1). In a Quantitative Linguistics fashion, the three corpora and the SMCP text collection are compared by means of non-parametric analyses of variance to compare distributions of linguistic variables and identify differences and similarities between them. The studied variables include vocabulary growth, word frequencies, lexical and key word densities, and grammar diversity as defined in papers I and II. Significant findings are quantified through the Probability of Superiority (PS) effect size, and the effects are ranked according to their magnitude. Finally, a "linguistic profile is created which effectively quantifies the observed language patterns as a coherent whole" (John, Brooks and Schriever 2017, p. 13).

Paper III leads to an effective method to discriminate idiosyncratic language patterns of Bridge Team Communication against other communicative settings (see figure 7). The sociolinguistic group of non-native speakers of English is profiled linguistically by building a "quantitative model of the language variety or genre of Bridge Team Communication as a sub-genre [...] of Maritime English" (ibid, p. 2).

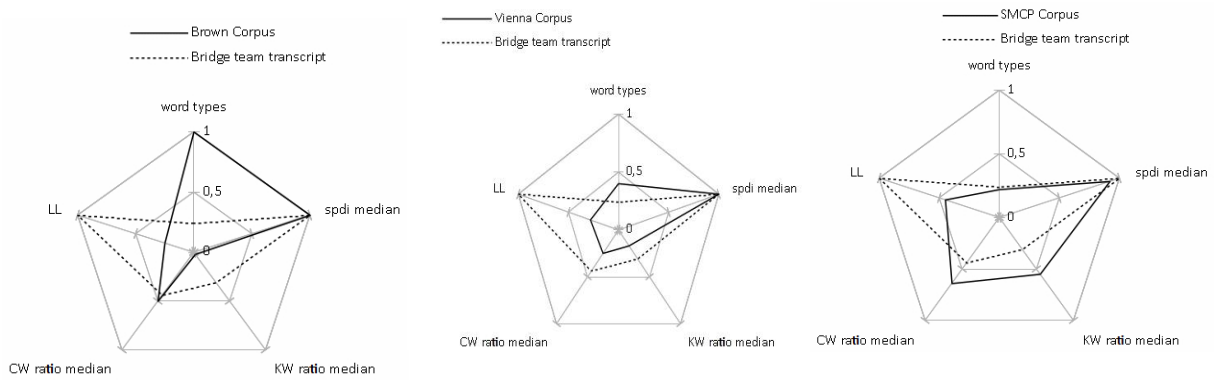

figure 7: Linguistic profile comparison of the analysed text corpora (John, Brooks and Schriever 2017, p. 12)

7.4. Results of paper IV

Paper IV titled "*Linguistic measurement of cognitive load in maritime team communication by native and non-native speakers of English*" pursues two objectives: Firstly, the two sociolinguistic groups of native and non-native speakers of English are compared by following the methodology applied in the previous paper. Secondly, proxy variables of cognitive load levels are analysed for each sociolinguistic group.

The first objective is approached by comparing the two groups' demographic meta-data as independent variables and their linguistics performance as the respective dependent variables. The comparison of the groups' demographics is performed to identify possible bias due to significant differences in the participants' independent variables (which include their gender, age, sea experience, etc.). A multiple regression analysis undertaken on the participants' meta-data highlights the similar composition of the sampled groups, and non-significant findings lead the researching student to disregard any impact on the groups' dependent linguistic variables.

Consequently, the language output by native and non-native speakers is compared in terms of different linguistic variables. This comparison follows the profiling method

introduced in paper III. However, while in the latter the sociolinguistic group of non-native speakers of English is compared with text corpora outside the maritime domain, this paper studies possible differences in the language patterns by native and non-native speakers involved in identical bridge team tasks.

The quantitative analyses of the different lexical and grammatical variables lead to very similar linguistic profiles for both groups. Small effect sizes between almost all dependent variables underline the applicability and robustness of the defined methods to discriminate Bridge Team Communication against other linguistic genres.

The similar findings for both groups also lead to the paper's second research question: To what extent does the cognitive load to produce similar linguistic structures differ in native and non-native speakers of English? To determine the cognitive load levels of the two sociolinguistic groups, proxy variables are studied by means of a Quantitative Content Analysis (QCA). This method singles out specific words which have been found to be uttered more frequently at high levels of cognitive load.

The undertaken Quantitative Content Analysis leads to significantly higher levels of cognitive load in non-native speakers than in native speakers. This is in line with the researching student's expectation. However, the degree to which cognitive load levels differ can hardly be quantified without the application of statistical methods.

Papers III and IV together prove that the sub-genre of Bridge Team Communication possesses a clearly defined linguistic profile. Native and non-native speakers of English produce specific linguistic patterns which are more closely related to each other than to verbal communication outside the maritime domain. This finding means that the applied methodology can indeed identify Bridge Team Communication without a qualitative analysis of its contents[61].

61 For a detailed analysis, see chapter *7.6. Overall results of research papers: creating a linguistic profile*.

7.5. Results of paper V

Paper V titled "*Speech acts in professional maritime discourse: A pragmatic risk analysis of bridge team communication directives and commissives in full-mission simulation exercises*" adopts a slightly different approach. While the Bridge Team Communication corpus is still used for data analysis, a more qualitative methodology has been adopted to identify risks of miscommunication. For this purpose, a Quantitative Content Analysis has been applied to extract speech acts including commissives and directives from the text corpus. These speech acts are then categorised qualitatively into possible risks of miscommunication. Finally, the findings are quantified by counting the instances of the different risk categories.

This mixed-methods approach has been adopted in order to provide a distribution of possible risks of miscommunication which can hardly be determined by applying a quantitative method alone. On the sender side of the Osgood & Schramm communication model, the risk analysis has been able to ascertain a substantial risk of communicative disruptions and misunderstandings due to discrepancies between the locutionary and illocutionary meaning of the messages. On the receiver's side of the model, a high percentage of the response utterances have also been found to provide ambiguous feedback to the original speaker. It was also found that the sociolinguist groups of native and non-native speakers produce similar risk communication frequencies caused by an insufficient use of disambiguation strategies (e.g. closed-loop communication).

7.6. Overall results of research papers: creating a linguistic profile

Each of the papers mentioned above studies Bridge Team Communication from a different perspective. By combining the findings of papers I to V, a linguistic profile can be created for making inferences on how closely related a given transcript of verbal communication is to the idiosyncratic Bridge Team Communication features identified in this research. In other words, profiling Bridge Team Communication results in *typical* statistical distributions of the analysed linguistic variables which can be used as benchmark values to determine if a given fragment of spontaneous verbal communication complies with these specific characteristics. The degree of this

compliance can then be used for assessing the appropriateness of the communication excerpt.

In the following section, the research questions listed in chapter *1.6. Research questions and hypotheses* are answered on the basis of the findings of the research papers summarised above. While the results of hypothesis testing, regression analyses and computation of the resulting effect sizes are included in the respective papers, the following sections aim to provide a more holistic explanation of the studied observations as an "explanatory sequential design" (Creswell 2014, p. 6).

Research question RQ_1 and its subordinate questions $RQ_{1.1}$ and $RQ_{1.2}$ provide the overarching objective of this research:

> RQ_1: How can idiosyncratic linguistic patterns of Bridge Team Communication be modelled by means of quantitative Corpus Linguistics methods?

Bridge Team Communication presents an idiosyncratic linguistic structure which can indeed be modelled by means of Quantitative Linguistics and Corpus Linguistics methods. Significant differences have been observed between the sampled Bridge Team Communication corpus and other linguistic settings outside the maritime domain. These are outlined in the following.

> $RQ_{1.1}$: Can the effectiveness of Bridge Team Communication be assessed by means of a specific lexical index measuring the exchange of maritime information?

On the one hand, lexical patterns differ strongly with regards to nautical key word frequencies. The chosen lexicon based on SMCP content words could be validated to provide an inherent statistical distribution which can be used as the expected (i.e. modelled) values in assessing naturalistic speech data. The key word distribution delivers a benchmark value computed on groups of native and non-native speakers of English. It has been found that the differences between the sampled L1 and L2 speakers are much smaller than the differences between Bridge Team Communication and non-maritime speech samples. As the sampled exercises are limited to standard navigational tasks, the resulting key word distribution may also be used to compare this benchmark with other communication content uttered by similar teams, e.g. to compare lexical structures in emergency situations or while participants suffer from fatigue, etc.

While a quantification of key word frequencies of nautical word types provides the most obvious results, the distribution of lexical density figures also leads to significant differences between Bridge Team and other verbal communication. In Bridge Team Communication, the proportion of content words to the total word count is closer to the lexical patterns of written language than to other naturalistic verbal exchange. Lexical density is known to express the degree to which information is shared between speakers. In the observed nautical setting, the primary objective for communicating verbally consists in exchanging navigational information. It is therefore no surprise that the lexical density distribution should include a high information content. Again, the reason of quantitative modelling is to provide a mathematical model to compare observed values against the expected distribution. The modelled values determined for dependent lexical variables provide a comparative framework for assessing transcribed spoken communication against a profiled benchmark value.

By combining the lexical density concept with key word frequencies, a specific index is capable of determining the degree of nautical idiomaticity while also delivering the extent to which nautical information is shared between speakers. Given that the exchange of unambiguous information is the paramount objective of Bridge Team Communication[62], lexical density and the developed specific index can be assumed to deliver a valid proxy for effective communication in this specific work-related discourse.

> $RQ_{1.2}$: Can the grammar diversity used by different speakers be compared by means of a special part-of-speech diversity index?

The index referred to in the answer to $RQ_{1.1}$ provides distributional information on the nautical information communicated between bridge team members. Given the high lexical density values, the sampled speech event leads to statistical figures which are much closer to written text than to spoken language. For this reason, an analysis of the speakers' grammar structure was required which would lead to a typical distribution for verbal communication. An index has been introduced which reuses the part-of-speech tags employed in the given lexical analysis. Computing this index provides a very efficient methodology to use existing information to ascertain if a given text fragment

62 For more details on the importance of unambiguous speech patterns in the maritime domain, see chapter *10.5 Paper V*, section *I. Introduction*.

belongs to the register of spoken language. A quantitative model has been introduced to discriminate efficiently between the two registers (i.e. written and spoken) and to provide an idiosyncratic structure of spontaneous verbal communication on a ship's bridge. This speech setting does not present the high level of grammatical diversity observed in the studied radio interviews but is very closely related to the speech samples collected in a text corpus of spoken language outside the maritime domain.

> RQ_2: Can quantitative Corpus Linguistic methods identify suitable patterns to assess the language performance and thus the effectiveness of Bridge Team Communication?

In language proficiency assessment, test takers typically find themselves in an environment where they can concentrate on the task at hand. The certified proficiency level is expected to warrant an effective communication in a real-life work environment on board ship where different tasks need to be attended to while communicating verbally with other members of the same bridge team. This research delivers a method to quantify typical and idiosyncratic linguistic distributions for the given speech event. The effectiveness of language performance is profiled by combining a set of dependent linguistic variables which have been found to deliver statistically significant differences. A quantitative model can be applied without the need of a qualitative, case-based assessment for which expert knowledge (by examiners, etc.) is required. Quantitative Corpus Linguistic models have thus been validated for identifying and assessing variables which are indicative of naturalistic language performance.

> $RQ_{2.1}$: To what extent do the speech patterns of bridge team communication by non-native speakers of English in full-mission simulation differ lexically and grammatically from other, non-nautical communication?

Bridge Team Communication differs sufficiently from other communicative settings to provide statistically significant distributions in linguistic variables for a series of linguistic features. Bridge Team Communication includes a higher information content than the other studied spoken text corpora while simultaneously making use of a relatively small number of word types. The chosen speech event centres around clearly defined navigational tasks which are discussed by team members. This leads to distinct idiosyncratic lexical patterns which are capable of identifying this particular speech

event against other communicative settings. Grammatically, Bridge Team Communication presents a very similar structure to spontaneous speech events outside the given communicative situation, albeit with a much higher information exchange and an idiosyncratic maritime lexicon.

> $RQ_{2.2}$: Can the exchange of technical information by bridge team members be assessed by means of Corpus Linguistics methods?

Corpus Linguistics methods reduce a possible bias observed in studies based on individual cases by reducing the impact of extreme values (e.g. outliers). The variables which have been studied on groups of native and non-native speakers of English result in a number of discrete values which have been modelled trough regression analyses and calculations of effect sizes. This approach results in a well-fitting quantitative model.

> $RQ_{2.3}$: Can nautical idiomaticity be assessed in authentic, inter-personal communication by bridge team members?

Nautical idiomaticity can be analysed by extracting content words from the prescriptive Standard Marine Communication Phrases (SMCP). Although these phrases have been produced on empirical findings to cover radio communication and a limited number of communicative settings on board ship, an analysis of word types included in the SMCP and the recorded full-mission exercises reveal that the SMCP key words used in this research provide a robust methodology to discriminate nautical from non-nautical communication.

> RQ_3: What differences can be observed in the linguistic patterns produced by bridge teams composed either of native speakers or of non-native speakers of English while performing identical navigational tasks?

Interestingly, the language genre of Bridge Team Communication leads to very similar distributional patterns in the dependent variables calculated on the groups of native and non-native speakers of English. Most observed between-groups differences in the computed lexical and grammatical frequencies result in non-significant findings or in effect sizes which are relatively low so that they can be assumed to result from sampling error alone. At the same time, a comparison of the two observed groups with authentic text genres outside a maritime settings validates the applicability of the adopted quantitative approach by delivering clear differences between the studied text corpora.

$RQ_{3.1}$: To what extent do the speech patterns of bridge team communication by non-native speakers of English differ lexically and grammatically from those produced by native speakers performing identical navigational tasks in full-mission simulation?

One salient lexical difference between L1 and L2 speakers is the ratio of word types to the produced word tokens. This so-called type-token ratio (TTR) is an indication of lexical variation in the speakers' utterances. Here, the observed native speakers make use of a wider mental lexicon than their non-native speaking counterparts.

$RQ_{3.2}$: In how far does the exchange of technical information by native speakers of English differ from that of non-native speakers?

While native speakers produce a more varied language, this is hardly reflected in the computation of their lexical density distribution. The difference between the native and non-native speaker group was found to be significant, albeit with a relatively low effect size which expresses the proximity of both distributional patterns. This can be explained by the increased use of synonyms by native speakers which leads to a higher type-token ratio but not to a higher lexical density. It can thus be argued that the exchange of information remains quite similar in both L1 and L2 speakers.

$RQ_{3.3}$: Can a difference in nautical idiomaticity by native and non-native speakers of English be observed in authentic, inter-personal communication by bridge team members?

The SMCP key word frequency calculated for the native and non-native speaker groups also results in a very low effect size which again indicates the similarity in the speech patterns produced by both groups. While contrasting nautical and non-nautical text corpora leads to highly significant differences in their dependent variables, the linguistic patterns uttered by the L1 and L2 teams are quite similar. This observation cross-validates the robustness of the adopted research approach as the analysed idiosyncratic features differ more strongly across the studied communicative settings than between speakers of different mother tongues.

However, while the sampled non-native speakers perform linguistically quite similarly, their observed cognitive load to produce similar language output was found to be significantly higher. To what extent the higher cognitive load levels can possibly have

negative consequences for a ship's navigation is not a part of this doctoral thesis. It does however identify an interesting area for further research.

> RQ_4: Can quantitative Corpus Linguistics methods assist in identifying risks of communicative disruptions and miscommunication by bridge team members?

Risks of communicative disruptions and miscommunication can be identified by means of Corpus Pragmatics methods. Possible risks can be identified in locutionary and illocutionary speech acts through a mixed-method approach aimed at isolating clues for discrepancies between the actual and the intended meaning of an utterance. Subsequently, the feedback given in the communicative loop can be analysed for the speakers' perlocutionary effect. Numerical values can be computed for discrete risk categories which provide distributional patterns. These patterns can then be analysed by means of hypothesis testing and effect size calculations.

> $RQ_{4.1}$: Can risks of miscommunication by bridge team members be identified by means of quantitative Corpus Linguistics methods?

Risks of miscommunication can be identified at various levels. The communicative loop (e.g. as represented by the Osgood & Schramm communication model) includes a speaker/sender and one or more listeners/receivers. From an information exchange perspective, communicative disruptions can occur when a receiver does not decode a message correctly. This misunderstanding can be caused by ambiguities in the speakers' utterances or an erroneous interpretation of the message by listeners. In the high-risk professional environment on a ship's bridge, all team members are expected to communicate in an unambiguous and precise manner. This research on authentic verbal exchange has been able to demonstrate that the sampled utterances sometimes differ quite starkly from the communicative clarity requested by the International Maritime Organization and other bodies.

> $RQ_{4.2}$: Can differences in the locutionary and illocutionary meaning be used to assess a potential risk of miscommunication by bridge team members?

Verbal exchange can be studied at different layers. While at surface level, the sampled native and non-native speakers of English deliver relatively similar dependent variables in their speech output, the ambiguity level in their utterances differs distinctly. Non-

native speakers tend to produce a higher number of possibly ambiguous speech acts due to an increased imprecision in the language used.

The adopted methodology has been capable of identifying differences between the locutionary and illocutionary meaning. While the actual effect of ambiguous utterances can only be measured by observing outcomes external to language production, an assessment of the inherent risk has been undertaken and validated. The aim of professional discourse is the clear and unambiguous exchange of information, and any risk of miscommunication can potentially have negative consequences. This research has contributed to an identification of these risks. The computed distribution can again be used as a benchmark value delivering a typical risk value. An improved risk awareness of exercise participants can be expected to be reflected in a reduced value of ambiguous messages.

> $RQ_{4.3}$: Can lexical structures in perlocutionary speech acts be used to make inferences on an actual miscommunication by bridge team members?

Lexical structures in perlocutionary speech acts can be analysed by means of Corpus Pragmatics methods by which discrete risk categories can be ascertained and quantified. By categorising the feedback from the communication partner into separate risk severity levels, the response of the listener to the original message delivers a clue as to the effectiveness of the given communicative loop. The decoding of the original message itself is a mental process which cannot be measured by linguistic variables. However, as the communicative loop is expected to be closed in Bridge Team Communication, the lack of this action can be counted, and so can incorrect or incoherent responses.

By answering the research questions raised in chapter *1.6. Research questions and hypotheses* a general overview has been given which is further detailed in the discussion and results sections of the respective research papers.

Profiling a linguistic behaviour consists in combining the variables described above. Each linguistic variable deals with one particular aspect of naturalistic verbal interaction. All linguistic variables together provide a set of idiosyncratic values for the given speech event of Bridge Team Communication sampled in standard navigational tasks. The profile hence delivers a set of statistical distributions which can be used as a reference or benchmark for the given communicative setting. The methodology allows

for an analysis of authentic communication without the need of an introspective evaluation. By developing the linguistic profile by means of Corpus Linguistics techniques on a special spoken corpus of authentic Bridge Team Communication by native and non-native speakers of English, the possible impact of extreme values is reduced due to regression to the mean effects. The method can be fully computerised so that transcripts of authentic verbal communication can be evaluated by comparing their distribution with the developed linguistic profile. Inferences can be made on individual or team performance which can be assessed without an expert-based evaluation of the linguistic performance.

8. Discussion and conclusions

This chapter summarises the implications of the adopted research methodology, it lists identified constraints and limitations and it provides some ideas for future research which takes into account the findings of the adopted approach. The chapter closes with final observations relating to the scope and applicability of the conducted quantitative profiling of Bridge Team Communication.

8.1. Implication of the research

This doctoral research set out to develop a quantitative methodology for a language performance assessment in the maritime domain. A quantitative assessment of authentic language performance can be automated, thus leading to a reduction of a case-based expert evaluation which is time-consuming and costly.

Training nautical officers in full-mission simulation exercises has shown that participants perform strikingly differently in a classroom language assessment and in authentic navigational exercises in which participants cannot concentrate on their language output but need to attend to a variety of other tasks simultaneously (Aarsæther & Moan 2007; Felsenstein, Benedict & Baldauf 2010; Hontvedt 2015). This difference between language *proficiency* and language *performance* has far-reaching consequences in a professional domain where “human erroneous actions” (European Maritime Safety Agency 2015, p. 8) causes four reported maritime incidents per day, on average. A fully computerised assessment of authentic language performance would lead to a more realistic evaluation of seafarers’ language competence in this very specific work area.

By profiling Bridge Team Communication in a quantitative manner, a case-based expert evaluation can be complemented in a mixed-methods fashion, but it also offers the potential for probabilistic assessment without any qualitative judgement of the subjects' communicative competence. However, assessing language performance without analysing its content can only be valid if it considers a range of linguistic variables which together provide a proxy for the overall construct of Bridge Team Communication. A linguistic profile is required as the benchmark against which speech samples are discriminated. This research has taken on the task of computing such a linguistic profile for the specific communicative environment of Bridge Team Communication.

8.2. Constraints and limitations of the research

A model may be considered a simplification of reality for the sake of clarity (McQuail and Windahl 1993). The objective of this doctoral research was to determine if a quantitative linguistic profile can be generated for the distinctive sociolinguistic communicative setting on a ship's bridge. For this purpose, a range of linguistic variables have been identified which study lexical and grammatical aspects of language on a semantic and pragmatic level. The selection of these variables is of course a simplification of the complex reality of any natural language. The modelled linguistic profile can thus only be regarded as a collection of quantitative linguistic variables which do not reflect the totality of natural languages beyond the scope of these variables. By confining the research to transcribed verbal communication, information conveyed by non-linguistic behaviour of bridge team members such as gestures, body language, etc. has been disregarded. Other aspects of naturalistic verbal exchange such as the speakers' intonation, pronunciation, pitch and others have also not been considered. The research has been undertaken from an information exchange perspective but has been limited to the information contained in the spoken utterances without considering other aspects of spoken language.

Another constraint is given by the analysed speech samples of Irish native speakers and German non-native speakers of English. This selection was carried out for pragmatic reasons. As stated in the introduction, hardly any contextualised transcripts of authentic Bridge Team Communication are available in the public domain, so that the researching

student was limited to a data collection within the time constraints of this doctoral research. Careful consideration has been given as to the generalisability of the research results beyond the sampled participant groups.

It is based on simulation sessions carried out with volunteering university students of Nautical Sciences. Here, a general symmetry in speech facilitates an ideal speech situation in which bridge team members participate under an assumption of equality (Habermas, 1979) which will not always be the case on board a sea-going ship.

Further, the validity of studying simulated versus authentic verbal communication needs to be considered. A qualitative comparison of the transcripts of real-life interaction analysed in Paper I with the exercises carried out in full-mission simulation reveals similar language patterns both lexically and grammatically. However, a quantitative analysis has not been undertaken due to the fact that no real-life pilotage operation has been included in the simulation exercises and would thus include a significant sampling error. Introspectively, the author would argue that full-mission simulation replicates the work environment on board ship in a much more authentic manner than a language proficiency assessment carried out in a classroom.

8.3. Future research

The validated methodology of profiling linguistic performance by team members has the scope to be extended beyond the confines of this doctoral research. With computer technology on the brink of being able to deliver automatic transcripts of spontaneous verbal communication, a much bigger database can be expected to be available in the future. This research needs be considered a first step towards a more comprehensive analysis of maritime and other communication based on naturalistic language performance data. The benchmark values provided here are not the end of this research, they rather constitute the beginning of a much wider research into a series of professional discourse settings. While language proficiency assessment will still be the most efficient way to evaluate linguistic competence, an evaluation of language performance provides a more realistic picture of how people communicate in a real professional context. This is of special importance in high-risk environments where miscommunication can have tragic consequences. This is not only true for the shipping industry but also for other professional domains where verbal communication is a key

competence. The observed discrepancy between language proficiency and communicative performance exists due to the fact that test takers can concentrate on one task (namely answering the assessment questions) while having to focus on several tasks simultaneously in a real-life communicative situation. A valid assessment methodology for naturalistic verbal interaction by bridge team members and in other professional settings should take into account this task simultaneity. This simultaneity is created in full-mission simulation exercises. However, in a purely classroom-based language assessment situation, further research in primary and secondary task settings is required.

8.4. Final observations

Bridge Team Communication can be considered a sub-genre of Maritime English which, as a linguistic genre, falls into the area of English for Specific Purposes (ESP). Bridge Team Communication refers to the exchange of professional information between navigating personnel on a ship's bridge. This verbal exchange of navigational information is trained intensively in full-mission ship-handling simulation facilities which authentically replicate the real work environment of deck officers. The training exercises engage trainees in a series of ship-handling tasks, one of which relates to sharing their mental picture of given situations with other bridge team members with the aim to develop a shared 'situational awareness', a concept that combines all aspects of cognition to build a shared understanding of the work tasks and context. The exchanges that occur in this highly relevant task for navigating a ship safely are undertaken by means of verbal communication, i.e. speaking and listening.

Participants of ship-handling simulation exercises need to listen and speak while they perform a series of other tasks simultaneously. The communicative context thus differs clearly from a language class or assessment in which participants can fully concentrate on language reception and production.

The linguistic and communicative performance in full-mission simulation is often assessed by one or more instructors. The instructors' assessment or feedback to participants mainly refers to the success of the navigational tasks at hand and giving advice on how to improve the technical aspects of navigation. Feedback concerning the linguistic and communicative performance is given as a qualitative statement (if at all).

In order to overcome this less than satisfactory situation the present doctoral research

has applied a number of measurements of linguistic variables with the intention to create an idiosyncratic linguistic profile. The linguistic profile is intended to identify statistically significant differences between different speakers which enable simulation instructors to give a more detailed feedback to simulation participants on their linguistic and communicative performance.

The attempt of profiling language performance has been a novel approach. For this reason, no similar studies have been available in the maritime context against which the measured variables could have been compared. This lack has motivated the researcher to record native-speaking participants in simulation exercises and use their linguistic profile as a control group for the non-native speaking participants.

Individual speaker performance does not only differ between the two socio-linguistic groups, it also varies considerably within the native speakers and the non-native speaker groups of English. To level out individual differences in speaker performance, a text corpus approach was adopted, with one representative sub-corpus for the recorded native speakers and another representative sub-corpus for the non-native speakers of English. The two sub-corpora thus include two socio-linguistic groups which differ in their native language while the demographic differences between both groups (e.g. the time they had been working at sea prior to the recordings) have not been found to present statistically significant differences.

The two socio-linguistics groups have been analysed for their use of lexical items and grammar structures. It is not surprising that the findings differ between native and non-native speakers of English. It was not the aim of this research to corroborate this anticipated finding, but rather to quantify these differences. Quantification leads to numerical values for the observed variables which allow statistical methods to be applied. This is a move away from an introspective assessment of linguistic performance of Bridge Team Communication by a simulation instructor.

The adopted approach has delivered a quantitative model for the professional communication structures in standard navigational situations on a ship's bridge. The model provides a benchmark against which communication in non-standard situations can be compared, for example in emergency situations. By linking the analysed linguistic variables to external factors, namely the given navigational situation, a better

understanding has been achieved as to the language performance by native and non-native speakers of English. Results indicate that the specific context of Bridge Team Communication leads to an idiosyncratic profile (or *footprint*). A deviation from this profile means that the employed language patterns lose their appropriateness which will increase the likelihood of communicative disrupts or plain misunderstanding. Conversely, complying with the linguistic profile means that communication is appropriate and fluent and allows crew members to share their 'mental model' (a mental representation of the surrounding world) verbally with the other team members and achieve a common understanding of a given situation. This increases the overall situational awareness of the whole bridge team.

The quantified lexical patterns and grammar structures provide a means for assessing the appropriateness of team communication and making inferences regarding the overall situational awareness of individual bridge team members and of the team as a whole. The truism all too often repeated in maritime accident reports that miscommunication and a lack of understanding by crew members is a contributory factor in shipping incidents underlines the importance in determining methods for inferring the psychological construct of shared mental models and others. Linguistic profiling is one possible method to approximate a construct which is not measurable directly. Given that the linguistic profiling method defined in this research delivers numerical values it can easily be combined with other psycho-cognitive methods for assessing psychological constructs.

In this research project, linguistic profiles have been created for the two socio-linguistic groups of native and non-native speakers of English. While clear differences have been determined in the language production by these two groups, some findings have brought about rather surprising results. The first of these is that the distribution of nautical keywords does not differ as strongly as may have been expected. Both socio-linguistic groups employ a quite restricted and relatively similar lexicon to share navigational information with other team members verbally. This is very much in line with the International Maritime Organisation's (IMO's) request for a clear and unambiguous maritime communication which seems to have found its way into the maritime education and training institutions of the studied countries. The assumption that native and non-native speakers communicate differently is commonplace. However, in the very specific

communicative context on a ship's bridge, lexical similarities prevail over the expected differences between both trainee groups.

This particular finding has motivated the researcher to look deeper into the language performance of the studied groups. After checking and verifying the specificity of the used measurement methods, a more pragmatic method (in the linguistic sense) has been adopted to study the language patterns beyond a surface level and analyse them with for their intended meaning. The unambiguity claim by the IMO (through coded language and closed loop communication structures) should ideally lead to an elimination of communicative disrupts or misunderstanding. When looking at differences between the trainees utterances and their intended meaning (i.e. the differences between their locutionary and illocutionary meaning), a high proportion of them have been found to be more ambiguous than the purely linguistic variables suggest. A high proportion of utterances still present a considerable risk of misinterpretation. The response to these utterances is often also not as unambiguous as one would hope. Here, the native speakers use a much wider variety of open communicative structures and adhere less to the coded language which the IMO advocates.

Today's crews on board sea-going ships are characterised by their multinational and multi-ethnic composition. This means that most team members speak English as a foreign language. Their language skills are typically certified by language proficiency tests which are assumed to measure the test takers' linguistic competence in a valid manner. However, their language performance may differ substantially when engaged in a socio-technical work environment in which multiple tasks need to be attended to simultaneously.

From an information exchange perspective, an insufficient language performance very probably leads to communicative disruptions and miscommunication. Bridge Team Members may utter ambiguous messages which are not interpreted correctly by other team members, and missing or incorrect feedback loops may lead to a distorted shared mental model being created and maintained. This discrepancy in the team member's mental models will increase the likelihood of erroneous decisions. In the high-risk environment of international shipping, miscommunication and a lack of situational awareness by one or more team members can have dramatic consequences causing

maritime accidents with injured or dead crew members, damages to ship and cargo and environmental pollution.

To avoid these scenarios, ship crews are trained in full-mission simulation exercises which truly replicate the work environment on a sea-going vessel. Training in correct or appropriate communication constitutes a major part of these bridge resource training exercises. An automated assessment of communication structures would support both the trainers and the trainees by providing them with empiric speech data. Being able to use realistic figures of the trainees' authentic linguistic performance would improve the learning process and reduce the costs and biases of a case-based assessment.

The author intends to contribute to this goal and to make the navigation of sea-going ships safer.

9. References

Adams, M.J., Tenney, Y.J. & Pew, R.W. (1995). Situation Awareness and the Cognitive Management of Complex Systems, *Human Factors, 37(1)*, 85-104.

Aarsæther, K.G., & Moan, T. (2007). Combined maneuvering analysis, AIS and full-mission simulation. TransNav, *International Journal on Marine Navigation and Safety of Sea Transportation, 1(1)*, 31-36.

Alderton, T. (2004). The global seafarer: Living and working conditions in a globalized industry. International Labour Organization. Geneva: International Labour Office.

Atkins, S., Clear, J., & Ostler, N. (1992). Corpus design criteria. *Literary and linguistic computing, 7(1)*, 1-16.

Baayen, R. H. (2001). Word frequency distributions. Dobrecht: Kluwer Academic Publishing.

Bachman, L.F. & Palmer, A.S. (1996). Language Testing in Practice. Oxford: Oxford University Press.

Bailey, N., Housley, W., & Belcher, P. (2006). Navigation, interaction and bridge team work. *The Sociological Review, 54(2),* 342-362.

Baker, P. (2006). Using corpora in discourse analysis. A&C Black.

Baker, P., Hardie, A., & McEnery, T. (2006). A glossary of corpus linguistics. Edinburgh: Edinburgh University Press.

Baker, P., & Egbert, J. (eds). (2016). Triangulating Methodological Approaches in Corpus Linguistic Research. New York: Routledge.

Baker, P. & Ellece, S. (2011). Key Terms in Discourse Analysis. London: Bloomsbury

Berns, M. & Matsuda, P.K. (2006). Applied Linguistics: Overview and history, Brown, K. (ed.), *The Encyclopedia of Language and Linguistics, 2nd edn*, Oxford: Elsevier, 394-405.

Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing 8(4)*, 243-257.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Boettger, R.K., & Palmer, L.A. (2010). Quantitative content analysis: Its use in technical communication. *IEEE Transactions on Professional Communication*, *53(4)*, 346-357.

Bocanegra-Valle, A. (2011). The language of seafaring: standardized conventions and discursive features in speech communications. *International Journal of English Studies*, *11(1)*, 35-53.

Bocanegra-Valle, A. (2010). Global markets, global challenges: the position of Maritime English in today's shipping industry. Linde López, A. & Crespo Jiménez, R.: *English in the European context: The EHEA challenge*, Bern: Lang, 151-174.

Bocanegra-Valle, A. (2013). Maritime English. C.A. Chapelle (ed). *The Encyclopedia of Applied Linguistics*. Malden: Wiley-Blackwell, 3579-3584.

Bod, R. (2003). Introduction to Elementary Probability Theory and Formal Stochastic Language Theory. Bod, R., Hay, J., & Jannedy, S. (Eds.). Probabilistic linguistics, 11-38. Massachusetts: Mit Press.

Bolstand, C.A. & Endsley, M.R. (2000). The effect of task load and shared displays on team situation awareness. *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society*.

Bowers, C. A., Jentsch, F., Salas, E., & Braun, C. C. (1998). Analyzing communication sequences for team training needs assessment. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 40(4)*, 672-679.

Brodje, A., Lundh, M., Jenvald, J., & Dahlman, J. (2013). Exploring non-technical miscommunication in vessel traffic service operation. *Cognition, technology & work, 15(3)*, 347-357.

Brown, J.D. (1988) Understanding research in second language learning: a teacher's guide to statistics and research design. Cambridge: Cambridge University Press.

Brown, J.D. (2004). Research Methods for Applied Linguistics: Scope, Characteristics,

and Standards, in: Davies, A. & Elder, C. (eds), *The Handbook of Applied Linguistics*, Malden: Blackwell.

Butler, J. (2013). Rethinking introspection: A pluralist approach to the first-person perspective. Berlin: Springer.

Cameron, D. (2001). Working with Spoken Discourse. London.

Canale, M. (1983), From communicative competence to communicative language pedagogy, in J. C. Richards and R. W. Schmidt (eds), *Language and Communication*. London: Longman, 2-27.

Canale, M. and Swain, M. (1980), Theoretical bases of communicative approaches to second language teaching and testing, *Applied Linguistics, 1*, 1-47.

Cannon-Bowers, J., Salas, E. & Converse, S. (1993). Shared Mental Models in Expert Team Decision Making. N. John Castellan (Ed.), Individual and Group Decision Making: Current Issues, 221-246. Hillsdale: Lawrence Erlbaum Associates.

Chauvin, C., Lardjane, S., Morel, G., Clostermann, J. P., & Langard, B. (2013). Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. *Accident Analysis & Prevention, 59*, 26-37.

Chomsky, N. (1965). Aspects of the Theory of Sntax. Massachusetts: MIT Press.

Chomsky, N. (1984). Modular Approaches to the Study of the Mind. San Diego: San Diego University Press.

Cole, C. & Trenkner, P. (2009). The Yardstick for Maritime English STCW assessment purposes. *IAMU Journal, 6 (1)*, 13-28. Tokyo: IAMU.

Cook, G. (2003). Applied Linguistics, Oxford University Press, Oxford.

Creswell, J.W. (2014). A concise introduction to mixed methods research. London: SAGE.

Crowdy, S. (1993). Spoken corpus design. *Literary and Linguistic Computing, 8(4)*, 259-265.

Dance, F.E.X. (1967). A helical model of communication. In: Dance, F.E.X. (ed). *Human Communication Theory*. New York.

Dance, F. E. X., & Larson, C. E. (1976). The functions of human communication: A theoretical approach. New York.

Davies, A. (2005). A Glossary of Applied Linguistics. Edinburgh.

Davies, A., & Elder, C. (eds). (2008). The handbook of applied linguistics. Malden: Blackwell.

DeFleur, M.L. (1970). Theories of Mass Communication. New York: McKay.

de la Campa Portela, R. (2005). Maritime casualties analysis as a tool to improve research about human factors on maritime environment. *Journal of Maritime Research*, *2(2)*, 3-18.

Demydenko, N. (2012). Teaching maritime English: A linguistic approach. *Journal of Shipping and Ocean Engineering, 2(4)*, 249-254.

Denzin, N.K. (1994). The Art and Politics of Interpretation. N.K. Denzin & Y.S. Lincoln (Eds.). *Handbook of Qualitative Research*, 500- 515. London: Sage.

Deutsch, K. W. (1966). The Nerves of Government: models of political communication and control; with a new introduction. New York.

Dörney, Z. (2007). Research Methods in Applied Linguistics. Oxford: Oxford University Press.

Dževerdanović-Pejović, M. (2013). Discourse of VHF Communication at Sea and the Intercultural Aspect. *International Journal for Traffic & Transport Engineering, 3(4)*, 377-396.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32-64.

Endsley, M. R. (1997). Supporting situation awareness in aviation systems. *Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation. 1997 IEEE International Conference on (Vol. 5)*, 4177-4181, IEEE.

Endsley, M.R. & Jones, W. (2013). Situation awareness. Lee, J.D. & Kirlik, A. (eds.) *The Oxford handbook of cognitive engineering*, *1*, 88-108. Oxford: Oxford University Press.

Entin, E. & Entin, E. (2000). Assessing team situation awareness in simulated military missions. Ergonomics for the new millennium. *Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society, Vol. 1*, 73–76.

European Commission (1999). The Marcom Project. The impact of multicultural and multilingual crews on maritime communication. Contract No WA-96-AM-1181. A Transport RTD Programme DG VII.

European Maritime Safety Agency (eds) (2015). Annual Overview of Marine Casualties and Incidents 2015, available at: http://www.emsa.europa.eu/emsa-documents/latest/item/2551-annual-overview-of-marine-casualties-and-incidents-2015.html

Faerch, C., & Kasper, G. (1987). Introspection in second language research (Vol. 30). Bristol: Multilingual Matters Limited.

Fan, L., Fei, J., Schriever, U., & Fan, S. (2017). The communicative competence of Chinese seafarers and their employability in the international maritime labour market. *Marine Policy, 83*, 137-145.

Felsenstein, C., Benedict, K., & Baldauf, M. (2010). Development of a Simulation Environment for Training and Research in Maritime Safety and Security. *Journal of Marine Technology and Environment, 3(2)*, 3-10.

Franceschi, D. (2014). The features of maritime English discourse. *International Journal of English Linguistics, 4(2)*, 78-87.

Frey, L., Botan, C. H., & Kreps, G. (2000). Investigating communication. New York.

Froholdt, L. L. (2010). Getting closer to context: a case study of communication between ship and shore in an emergency situation. *Text & Talk – An Interdisciplinary Journal of Language, Discourse & Communication Studies, 30(4)*, 385-402.

Fulcher, G. & Davidson, F. (2012). The Routledge Handbook of Language Testing. New York: Routledge.

Garbis, C. & Artman, H. (1998). Team communication and coordination as distributed cognition, *9th Conference of Cognitive Ergonomics*, 151-156.

Garside, R. (1987). The CLAWS Word-tagging System. in: R. Garside, G. Leech and G. Sampson (eds), *The Computational Analysis of English: A Corpus-based Approach.* London: Longman.

Garside, R. (1996). The robust tagging of unrestricted text: the BNC experience. J. Thomas and M. Short (eds). *Using corpora for language research: Studies in the Honour of Geoffrey Leech Longman*, London: Longman, 167-180.

Garside, R., Smith, N. (1997) A hybrid grammatical tagger: CLAWS4, in: Garside, R., Leech, G., and McEnery, A. (eds.) *Corpus Annotation: Linguistic Information from Computer Text Corpora.* London: Longman, 102-121.

Gerbner, G. (1967). Mass media and human communication theory. Dance, F.E.X. (Ed). *Human Communication Theory*. New York: Holt, Rinehart and Winston, 40-60.

Gramley, S. (2008). English for Specific Purposes (ESP). In: Gramley, S. & Gramley, V. (eds). Bielefeld Introduction to Applied Linguistics: A Course Book. Bielefeld.

Guba E.G. & Lincoln, Y.S. (1994). Competing Paradigms in Qualitative Research. Denzin, N.K. & Y.S. Lincoln. Handbook of Qualitative Research, 105-117. London: Sage.

Hontvedt, M. (2015). Simulations in maritime training: A video study of the socio-technical organisation of ship simulator training, Doctoral Thesis, University of Oslo.

Lincoln, Y.S., Lynham, S.A., & Guba, E.G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. N.K. Denzin & Y.S. Lincoln (Eds.). *The SAGE Handbook of Qualitative Research*, *4*, 97-128, London: Sage.

Haig, K. M., Sutton, S., & Whittington, J. (2006). SBAR: a shared mental model for improving communication between clinicians. *The joint commission journal on quality and patient safety*, *32*(3), 167-175.

Hall, C.J., Smith, P.H. & Wicaksono, R. (2011). Mapping Applied Linguistics: A Guide for Students and Practitioners, London: Routledge.

Hammersley, M. (1987). Some notes on the terms 'validity' and 'reliability. *British Educational Research Journal, 13 (1)*, 73-81.

Hardt-Mautner, G. (1995). Only Connect: Critical Discourse Analysis and Corpus Linguistics, UCREL Technical Paper 6. Lancaster: University of Lancaster.

Harris, Z. (1952), Discourse analysis, *Language, 28*. 1-30.

Hatch, E. & Lazaraton, A. (1991). The Research Manual: Design and Statistics for Applied Linguistics, Rowley: Newbury House.

Heaton, J.B. (1975). Writing Language Tests. London: Longman.

Hetherington, C., Flin, R., & Mearns, K. (2006). Safety in shipping: The human element. *Journal of safety research, 37(4)*, 401-411.

Hontvedt, M., & Arnseth, H. C. (2013). On the bridge to learn: Analysing the social organization of nautical instruction in a ship simulator. *International Journal of Computer-Supported Collaborative Learning, 8(1)*, 89-112.

Horck, J. (2005). Getting the best from multi-cultural crews. Paper presented at BIMCO 100 years and GA 2005, Copenhagen.

Hymes, D. (1972). On communicative competence, in J. B. Pride and J. Holmes (eds), *Sociolinguistics: Selected Readings*. London: Penguin, 269-93.

Hymes, D. H. (1986). Discourse: Scope without depth. *International Journal of the Sociology of Language*, 57. 49-89.

International Chamber of Shipping (eds.). (2015). Manpower Report 2015. London: Marisec Publications.

International Maritime Organization. (1977). Convention on the International Regulations for Preventing Collisions at Sea (COLREGs). London: IMO.

International Maritime Organization. (1978). Standard Marine Navigational Vocabulary (SMNV). London: IMO.

International Maritime Organization. (1978). International Convention on Standards of Training, Certification and Watchkeeping for Seafarers. London: IMO.

International Maritime Organization. (2001). Standard Marine Communication Phrases. London: IMO.

International Maritime Organization. (2010). The Manila Amendments to the Seafarers' Training, Certification and Watchkeeping (STCW) Code. London: IMO.

International Maritime Organization. (2014). International Convention for the Safety of Life at Sea (SOLAS). Consolidated edition 2014. London: IMO.

International Maritime Organization. (2015). Model Course 3.17 "Maritime English". London: IMO.

Iordanoaia, F. (2010). Master of the ship, manager and instructor. *Management & Marketing Journal, 1*, 133-155.

Jensen, R.S. (1997). The Boundaries of Aviation Psychology, Human Factors, Aeronautical Decision Making, Situation Awareness, and Crew Resource Management, *International Journal of Aviation Psychology, 7(4)*, 259-267.

John, P., Björkroth, P. & Noble, A. (2013). Making SMCP count! *Proceedings of the International Maritime English Conference (IMEC25)*, Istanbul, Turkey, 136-151.

John, P., Noble, A., & Björkroth, P. (2016). Low-fi simulation of bridge team communication. *WMU Journal of Maritime Affairs, 15(2)*, 337-351.

Johnson, B. (1999). English in the Global Maritime Distress and Safety System. *World Englishes. 18(2)*. 145-157.

Johnson, F.C. & Klare, G.R. (1961). General models of communication research: A survey of the developments of a decade. *Journal of Communication, 11(1)*, 13-26.

Johnston, J.M & Pennypacker, H.S. (1980). *Strategies and tactics of human behavioural research*. Hillsdale: Lawrence Erlbaum Associates.

Jones, N. (20012). Reliability and dependability. In: G. Fulcher & F. Davidson (Eds.).*The Routledge handbook of language testing,* 350-362. New York: Routledge.

Jurkovič, V. (2015). Shared and Specific Features of Maritime English within the LSP Context. *Facetten der Fachsprachenvermittlung Englisch – Hands on ESP Teaching, 5,* Berlin, 185-205.

Kataria, A. (2011). Maritime English and the VTS. *Proceedings of the International Maritime English Conference (IMEC23)*, Constanta, Romania, 25-33.

Kataria, A. (2015). An ethnographic exploration of ship-shore communication. *(Doctoral dissertation, Cardiff University)*.

Kataria, A., Holder, E., Praetorius, G., Baldauf, M., & Schröder-Hinrichs, J.-U. (2015). Exploring Bridge-Engine Control Room Collaborative Team Communication. *TransNav - the International Journal on Marine Navigation and Safety of Sea Transportation, 9(2)*, 169-176.

Koplenig, A. (2017). Against statistical significance testing in corpus linguistics. *Corpus Linguistics and Linguistic Theory,* 1-26.

Lane, A.D., Kahveci, E. & Sampon, H. (2002). The Formation and Maintenance of Transnational Seafarer Communities. Cardiff: Seafarers International Research Centre.

Lasswell, H.D. (1948). The structure and function of communication in society. Bryson (ed.). (1948). *The Communication of Ideas*. New York.

Layder, D. (1993). New Strategies in Social Research. Cambridge: Polity Press.

LeCompte, M.D., & Schensul, J.J. (2010). Designing and conducting ethnographic research: An introduction (Vol. 1). Lanham: Rowman & Littlefield.

Leech, G. (1991). The state of the art in corpus linguistics, in: *K. Aijmer & B. Altenberg (eds) English Corpus Linguistics: Studies in Honour of Jan Svartvik. London:* Longman, 105-122.

Leech, G. (1992). Corpora and theories of linguistic performance. Svartvik, J. (Ed.) *Directions in corpus linguistics. Proceedings of Nobel Symposium 82*, Stockholm, Sweden, 105-122. Berlin: Mouton de Gruyter.

Leech, G. (1997). Introducing Corpus Annotation, R. Garside, G. Leech & A. McEnery (Eds.). *Corpus Annotation*, 1-18. London: Longman.

Leech, G., Garside, R., Bryant, M. (1994). CLAWS4: The tagging of the British National Corpus. *Proceedings of the 15th International Conference on Computational Linguistics (COLING 94) Kyoto, Japan*, 622-628.

Lees, G., & Williamson, W. (2013). Handbook for Marine Radio Communication 5E. London: Taylor & Francis.

Lloyd's List Intelligence (eds) (2016): Lloyd's List Intelligence Casualty Statistics, available at: http://www.lloydslistintelligence.com/llint/home-casualties/index.htm, as of 26/01/2015.

Marine Accident Investigation Branch (2010). Report on the investigation of the grounding of MV Maersk Kendal on Monggok Sebarok reef in the Singapore Strait on 16 September 2009, London. Available at: https://assets.digital.cabinet-office.gov.uk/media/547c6fc9ed915d4c10000041/MaerskKendalReport.pdf as of 28/04/2016.

Marine Board (ed). (1994). Minding the Helm: Marine Navigation and Piloting. Washington: National Academies Press.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E. & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, *85*(2), 273-283.

Mayo, D. G., & Cox, D. R. (2006). Frequentist statistics as a theory of inductive inference. *Lecture Notes-Monograph Series*, 77-97.

McCarthy, M. (2001). Issues in Applied Linguistics. Cambridge: Cambridge University Press.

McEnery, T., & Wilson, A. (2001). Corpus linguistics: An introduction. Edinburgh: Edinburgh University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). Corpus-based language studies: An advanced resource book. London: Taylor & Francis.

McNamara, T. (2000). Language Testing. Oxford: Oxford University Press.

McNeill, P. (1990). Research Methods. Second Edition. New York: Routledge.

McQuail, D., & Windahl, S. (1993). Communication models for the study of mass communications. Second edition. New York: Routledge.

Mokros, H. B., & Deetz, S. (1996). What counts as real?: A constitutive view of communication and the disenfranchised in the context of health. E.B. Ray (Ed.). *Communication and disenfranchisement: Social health issues and implications*. 29-44.

Molt, E. (2006). No Double-Dutch at Sea: How English Became the Maritime Lingua Franca. *International Journal of Maritime History, 18(2)*. 245-255.

Nakamura, J. & Sinclair, J. (1995) The world of woman in the Bank of English: internal criteria for the classification of corpora, *Literary and Linguistic Computing 10(2),* 99-110.

National Transportation Safety Board (2008a). Accident report NTSB/MAR-08/01PB2008-916401 Heeling Accident on M/V Crown Princess Atlantic Ocean Off Port Canaveral, Florida, Washington. Available at: http://www.ntsb.gov/investigations/AccidentReports/Reports/MAR0801.pdf as of 28/04/2016.

National Transportation Safety Board (2008b). Accident report NTSB/MAR-08/02PB2008-916402 Grounding of U.S. Passenger Vessel Empress of the North, Washington. Available at: http://www.ntsb.gov/investigations/AccidentReports/Reports/MAR0802.pdf as of 28/04/2016.

National Transportation Safety Board (2009). Accident report NTSB/MAR-09/01PB2009-916401, Washington. Available at: http://www.ntsb.gov/investigations/AccidentReports/Reports/MAR0901.pdf as of 28/04/2016.

Newby, H. (1977). In the field: reflections on the study of Suffolk farm workers. C. Bell and H. Newby (Eds). *Doing Sociological Research.* London: Allen & Unwin.

Niesser, U. (1976), Cognition and Reality: Principles and Implications of Cognitive Psychology. San Francisco: Freeman.

Noble, A. (2017). Maritime English put to the test! The feasibility and desirability of setting global standards for Maritime English: a survey-based study. Doctoral thesis, University of Antwerp, Belgium.

Noble. A, Björkorth, P. & John, P. (2014). Exploiting the didactic possibilities of low-fi simulation in virtual bridge team communication exercises, *Proceedings of the International Maritime English Conference*, Terschelling, The Netherlands, 159-174.

Noble, A., Vangehuchten, L., & van Parys, W. (2011a). Intercultural Competence and

Effective Communication at Sea: An Invitation to Celebrate Diversity on Board. *Proceedings of the International Maritime English Conference*, Constanta, Romania, 131-149.

Noble, A., Vangehuchten, L., & van Parys, W. (2011b). Communication for maritime purposes: some exploratory results of a survey-based study on intercultural and linguistic features. *ITL. International journal of applied linguistics, (162)*, 111-133.

Novi, C. (1999) Multilingual Harmonization and Standardization of Technical Terminology an the International Maritime Organization. D, Newman & M. Van Campenhoudt: Maritime Terminology, Brussels,:Les Editions du Hazard, 110-125.

Nunan, D. (1992). Research Methods in Language Learning, Cambridge University Presse, Cambridge.

Øvergård, K. I., Nielsen, A. R., Nazir, S., & Sorensen, L. J. (2015). Assessing navigational teamwork through the situational correctness and relevance of communication. *Procedia Manufacturing, 3*, 2589-2596.

Orasanu, J. (1990). Shared mental models and crew decision making. *12th Annual Conference of the Cognitive Science Society*.

Paltridge, B. (2006). Discourse Analysis: An Introduction. London: Continuum.

Partington, A. (2003). The Linguistics of Political Argument: The Spin-doctor and the Wolf-pack at the White House. New York: Routledge.

Pawlak, M. & Aronin, L. (2014). Essential Topics in Applied Linguistics and Multilingualism: Studies in Honor of David Singleton. New York: Springer.

Pearce, W. B. (1995). Public dialogue & democracy: A guide for the discussion leader. Chicago.

Pennycook, A. (2001). Critical Applied Linguistics: A Critical Introduction, London: Lawrence Erlbaum Associates.

Phakiti, A. & Paltridge, B. (2015). Approaches and Methods in Applied Linguistics Research, A. Phakiti & B. Paltridge. *Research Methods in Applied Linguistics: A Practical Resource*. London: Bloomsbury Publishing.

Porathe, T., Eklund, P., & Goransson, H. (2014). Voice and Text Messaging in Ship Communication. Stanton, N., Landry, S., Di Bucchianico, G. & Vallicelli, A. (eds). *Advances in Human Aspects of Transportation: Part I*, 183-191.

Pritchard, B. (2002). On the Standards of Maritime English - Pedagogical implications. *Proceedings of the Seminar on Maritime English – Istanbul 2002*, Istanbul Technical University & JICA, 68-81.

Pritchard, B. (2003). Maritime English syllabus for the modern seafarer: Safety-related or comprehensive courses?. *WMU Journal of Maritime Affairs, 2(2)*, 149-166.

Pritchard, B. & Kalogjera, D. (2000). On Some Features of Conversation in Maritime VHF Communication. In M. Coulthard, J. Cotterill & F. Rock (Eds.). *Dialogue Analysis VII: Working with Dialogue: Selected Papers from the 7th IADA Conference Birmingham 1999,* 185–194.

Pyne, R., & Koester, T. (2005). Methods and means for analysis of crew communication in the maritime domain. *The Archives of Transport, 17(3-4)*, 193-208.

Riff, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research*. New York: Routledge.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, *52(1)*, 5-18.

Salas, E., Prince, C., Baker, P.D. & Shrestha, L. (1995). Situation awareness in team performance. *Human Factors, 37*(1), 123–126.

Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, *9*(4), 297-323.

Sampson, H. A. (2013). International seafarers and transnationalism in the twenty-first century. Manchester: Manchester University Press.

Sampson, H., & Zhao, M. (2003). Multilingual crews: communication and the operation of ships. *World Englishes, 22(1)*, 31-43.

Schmitt, N. (2002). An Introduction to Applied Linguistics, London: Arnold.

Schramm, W. (1954). The Process and Effects of Mass Communication. Champaign: University of Illinois Press.

Schriever, U.G. (2005). A discourse on teaching and learning of maritime English in the context of different linguistic and cultural backgrounds, *Proceedings of the 17th International Maritime English Conference, 4-7 October, Marseille, France*, 1-14.

Schriever, U.G. (2008). Maritime communication in an international and intercultural discourse. Doctoral thesis, The University of Tasmania.

Sharma, A., & Nazir, S. (2017). Distributed Situation Awareness in pilotage operations: implications and challenges. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation, 11*, 289-293.

Sinclair, J. M. and Coulthard, M. (1975), Towards an Analysis of Discourse, Oxford.

Smith, K. & Hancock, P.A. (1995) Situation awareness is adaptive, externally directed consciousness. *Human Factors*, *37*, 137–148.

Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors*, *41*(1), 61-71.

Strevens, P. (1983) A Case-History in the Constructions of International Maritime English, Illinois: University of Illinois.

Strevens, P.; Weeks, F. (1985) The Creation of Regularised Subset of English for Mandatory Use in Maritime Communications: *SEASPEAK, Language Planning Newsletter, Vol. 11(2)*, Honolulu, 1-5.

Stubbs, M. (2008). Language Corpora, in: Davies, A. & Elder, C. (eds), The Handbook of Applied Linguistics, Hoboken: Blackwell Publishing.

Swales, J. (1986). Citation analysis and discourse analysis. *Applied Linguistics*, 7(1), 39-56.

Swales, J. M. (1990). Discourse analysis in professional contexts. *Annual Review of Applied Linguistics*, 11, 103-114.

Tognini-Bonelli, E. (2001). The corpus-driven approach, W. Teubert and R. Krishnamurthy (Eds.). *Corpus Linguistics at Work*, Amsterdam: John Benjamins.

Trenkner, P. (1996). IMO-Standard Marine Communication Phrases (SMCP)–an attempt to meet increased communication requirements of ship's officers. *Ninth International Maritime Lecturers' Association International Conference on Maritime Education and Training (IMLA 9)*, Kobe, Japan.

Trenkner, P. (2000). Maritime English. An attempt of an imperfect defintion. *Proceedings of the Second IMLA Workshop on Maritime English in Asia (WOME 2A)*, 1-8.

Trenkner, P. (2002). The IMO Standard Marine Communication Phrases (SMCP) and the requirements of STCW Convention 1978/95. *Proceedings of International Seminar on Maritime English*, Istanbul, Turkey, 20-22.

Trenkner, P., & Cole, C. (2010). Raising the Maritime English bar: The STCW manila Amendments and their impact on Maritime English. *Proceedings of the International Maritime English Conference*, Alexandria, 3-16.

Vasishth, S., & Nicenboim, B. (2016). Statistical methods for linguistic research: Foundational ideas–Part I. *Language and Linguistics Compass*, *10(8)*, 349-369.

Wagenmakers, E. J., Lee, M., Lodewyckx, T., & Iverson, G. J. (2008). Bayesian versus frequentist inference. *Bayesian evaluation of informative hypotheses*, 181-207.

Weintrit, A. (ed). (2009). Marine navigation and safety of sea transportation: Advances in Marine Navigation. Boca Raton: CRC Press.

Weeks, F., Glover, A.,Johnson, E. & Strevens, P. (1984) SEASPEAK – Reference Manual Manual. Oxford: Pergamon Press.

Weir, C.J. (2005). Language Testing and Validation. New York: Palgrave Macmillan.

Winbow, A. (2002, March). The importance of effective communication. *International Seminar on Maritime English*, Istanbul: IMLA, 20-22.

Wittgenstein, L. (1953). 1958. *Philosophical investigations (Manuscript 1945-1949.* Translated from Philosophische Untersuchungen by GEM Anscombe.

10. Appended papers

10.1. Paper I

Information density in bridge team communication and miscommunication – a quantitative approach to evaluate maritime communication[1]

Peter John

Faculty of Maritime Studies, Jade University of Applied Sciences, Germany

Australian Maritime College, University of Tasmania, Australia

Benjamin Brooks

Australian Maritime College, University of Tasmania, Australia

Christoph Wand

Faculty of Maritime Studies, Jade University of Applied Sciences, Germany

Ulf Schriever

Australian Maritime College, University of Tasmania, Australia

Abstract – The paper presents a method for measuring the information flow in bridge team communication based on an analysis of the voyage data recorder transcript of MV *Cosco Busan* alliding with a bridge pillar in San Francisco Bay in dense fog. The aim is to assess the efficiency of maritime communication by developing a specific index suitable for evaluating individual and team communication performance. The allision investigated in this paper leads to the conclusion that the information flow in bridge team communication can be measured quantitatively and that a correlation to external situational factors can be demonstrated.

Key words – *Maritime communication, Bridge team communication, Voyage data recorder, Quantitative linguistics*

1 The final publication is available at Springer via http://dx.doi.org/10.1007/s13437-013-0043-8.

1. Introduction

The importance of an effective and successful bridge team communication has been widely acknowledged by the maritime industry and the International Maritime Organisation (IMO) as a decisive factor to ensure the safe operation of any sea-going vessel. A large number of marine accidents involve communication errors which lead to erroneous situational awareness and flawed decisions by the officers conning the ships. A notorious example is the collision between MV *Lykes Voyager* and MV *Washington Senator* in April 2005 due to a VHF communication error (Marine Accident Investigation Branch 2005).

In an effort to minimise misunderstandings, the IMO included communication skills in its Standards of Training, Certification and Watchkeeping (STCW) 1978 Convention and in its 1995 amendments (International Maritime Organization 1978; 2010). The Standard Marine Communication Phrases (SMCP) were adopted and made mandatory by the IMO resolution A.918 (22) in November 2001:

> 'As navigational and safety communications from ship to shore and vice versa, from ship to ship, and on board ship must be precise, simple and unambiguous so as to avoid confusion and error, there is a need to standardize the language used. This is of particular importance in the light of the increasing number of internationally trading vessels with crews speaking many different languages, since problems of communication may cause misunderstandings leading to dangers to the vessel, the people on board and the environment.'
>
> (International Maritime Organization 2001, p. 3)

In 2010, the Manila Diplomatic Conference on the STCW Convention also approved a number of significant amendments including a stronger focus on an effective oral communication in accordance with chapter V, regulation 14, paragraphs 3 and 4 of the SOLAS Convention (International Maritime Organization 2004).

Although communication is recognised as key to safe navigation, most research activities have either focussed on isolated marine incidents or have been part of a wider research conducted on human factors, for example, where communication has been identified as a contributing factor to ship safety.

This paper aims to deliver a quantitative method for assessing maritime communication

and the information shared by the communicating parties. By applying this method, *ineffective* communication can be detected and the analysis' results be used for training active and future seafarers.

2. Literature review

The implications of the definitions by the STCW 1978 Convention and its 1995 and 2010 Manila amendments have been discussed at the International Maritime English Conference (Cole 1994; Trenkner and Nielsen 1998; Trenkner 2002; Cole and Trenkner 2009; Trenkner and Cole 2010; Cole and Trenkner 2012).

Research studies of past decades identified communication problems to be responsible for up to 40% of marine casualties (Quinn and Scott 1982; Marine Transportation Research 1981; Paramore et al. 1979; UK Department of Transport 1991). The Marcom Project studied maritime communication by multicultural and multilingual crews and found that 40% of human errors were 'communicatively relevant factors' (The Marcom Project 1999).

More recent research on maritime incidents in which communication was a decisive factor was conducted by McCallum, McCrae and Pyne and Koester: McCallum et al. studied the contribution of 'ineffective communications to marine casualties' developing a 'standardized procedure [...] to guide casualty investigations by Coast Guard investigators over a six-month period' (McCallum et al. 2000, p. 384) and which identified, amongst others, the following specific problems: 'did not communicate [...]; did not question others to assert own interpretation [...]; did not verify information validity or accuracy [...]' (McCallum et al. 2000, p. 389).

According to research carried out by McCrae, a MAIT analysis 'revealed that 41.2 and 23.5% of situational factors in groundings were accounted for by poor communication and an undermanned bridge respectively' (McCrae 2009, p. 29).

Pyne and Koester analysed a 'number of maritime accident reports in which a failure of effective crew communication played a central role in the causal chain' (Pyne and Koester 2005, p. 1). They stated that '[t]he need for clear verbal communications between parties in the commercial marine environment is multi faceted as the ship is the working environment, learning environment and social environment for its personnel.

[...] [B]oth Pilot and crew must be able to communicate effectively to ensure safety.' (Pyne and Koester 2005, p. 7)

Other researchers have used taxonomies to carry out qualitative communication analysis from a psychological perspective by applying concepts of human reliability as a probable cause of error (Chauvin 2011), discursive psychology, principles of conversation analysis, Wittgenstein's philosophy (Froholdt 2010) and situational awareness (Grech 2005) to a maritime setting. Chauvin analysed 'key models and concepts developed in psychology which help to understand the role of human factors in complex systems' (Chauvin 2011, p. 625). Referring to the collision between MV *Cuyahoga* and MV *Santa Cruz* she states that: 'Dialogue between the three members of the crew would have certainly helped avoid the accident. But each thought, wrongly so, that they shared the same representation of the situation' (Chauvin 2011, p. 628). Froholt examined 'how non-native speakers orient toward and manage potential human error' (Froholt, p. 385). In an analysis of a telephone call between a shipmaster and a shipping company's superintendent both of which were non-native English speakers she arrived at the conclusion that they used 'discursive remembering and emotion displays, and they accomplished a collaborative understanding' (Froholt, p. 397). Grech stated that a 'lack of situation awareness occurred because of failure of communication between crewmembers, especially when relevant pieces of important information were not communicated' (Grech 2005, p. 83).

In his book 'Cognition in the wild,' Edwin Hutchins modelled 'the communication between networks as external inputs applied directly to the units in each network' (Hutchins 1995, p. 250). He arrived at some interesting conclusions: 'Our simulations provide us with a means to answer this question. They indicate that more communication is not always in principle better than less. Under some conditions, increasing the richness of communication may result in undesirable properties at the group level' (Hutchins 1995, p. 252). 'Thus, diversity of interpretations is fairly easy to produce as long as the communication among the members of the community is not too rich' (Hutchins 1995, p. 255).

Apart from the maritime area, communication has also been studied in other high-risk environments like aviation (Jentsch et al. 1999; Salas et al. 1999; Hobbs and Williamson 2003; Salas et al. 2006) and nuclear power plants (Kim et al. 2011; Rodrigues de

Carvalho et al. 2012) or where communication is an essential part of team work (Bowers et al. 1998; Koiso and Nishida 1999; Cooke et al. 2000; Chow et al. 2000; Chin et al. 2002; Gonzalez 2005; Levin 2006; Webster and Cao 2006; Mishra and Mishra 2009; Khawaja et al. 2012).

In an effort to help prevent marine accidents by informing the maritime community, the British Marine Accident Investigation Branch publishes annual reports which include recommendations to 'individual organisations, owners or companies that are specific to their area, vessel or company. [...] [T]he lessons are promulgated through the individual investigation reports' (Marine Accident Investigation Branch 2009, p. 31). Accepted recommendations to improve maritime communication include routine tasks (e.g. in cargo operations) and emergency actions carried out by crew members, as well as face-to-face and radio communication with tug and mooring parties, shore management, ISM auditors and other vessels. Examples are given in Fig. 1 (MAIB recommendations). These reports recommend improvements to ensure an effective communication on board. However, no clear definition is given as to when communication can actually be considered effective.

Report no.	Investigation	Recommendation
2007-176	Annabella	"[…] to ensure that: effective communications and procedures exist […] to ensure ship's staff have the resources and opportunity to safely oversee the loading and securing of cargo." ([Marine Accident Investigation Branch, 2007], p17)
2007-185	Sea Express 1 / Alaska Rainbow	"Review your Safety Management System with particular respect to: […] improving external communications in the event of an emergency […]". (ibid., p41)
2007-191	Velazquez	"Review and revise your company operating procedures to ensure: […] Effective communications are maintained between tug crews and ship mooring parties […]" (ibid., p42)
2008-168	Sichem	"Ensure that only the agreed working language is spoken in

	Melbourne	work related to communications." ([Marine Accident Investigation Branch, 2008], p20)
2009-110	Moondance	"Provide guidance to suitably trained internal ISM auditors on the scope of their responsibilities, including […] communications." ([Marine Accident Investigation Branch, 2009], p36)
2009-113	RIB 6	"[…] to ensure that all centre activities are effectively supported by appropriate means of communication at all times." (ibid., p36)
2009-130	Maersk Newport	"Review and amend your current procedures to ensure: […] Effective and inclusive communications between shore management, contractors and ship's staff [...]" (ibid., p38)

Fig 1: MAIB recommendations

3. Research methodology and results of analysis

The publications cited in the previous chapter refer to concepts such as *accurate, clear, better, effective, efficient, inclusive, multi-faceted, rich* or *valid* communication. They all assume that there is a common understanding of what is actually meant by the terms *effective, valid* and so on.

However, an extensive, empirical research to measure and assess the efficiency or validity of maritime communication has not been conducted and no empirical tool has been developed to estimate the information flow of bridge team communication in a systematic manner.

This article proposes a quantitative methodology to measure the information content in naturally occurring speech by the different members of a bridge team, including radio communication. Through the use of computational linguistics, the density of information is extracted for different conversation segments and related to the events taking place on board the ship.

For the purpose of this research, the term bridge team conversation (or BTC) shall refer to all spoken utterances by any bridge team member, either face-to-face or by radio

(VHF and UHF). The main function of BTC is the exchange of information in order to increase the situational awareness of the communicators and eventually contribute to a safe navigation. According to Endsley, situational awareness is 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future' (Endsley 1995, p. 36).

Other important factors inherent to any conversation include social, psychological and other aspects. However, the main purpose of communication within a bridge team environment is the exchange of information in order to ensure a safe navigation. This is why these other factors are deemed of secondary importance and are disregarded in this paper.

The sample used for presenting the method is the transcription of the Voyage Data Recorder on board MV *Cosco Busan* which in November 2007 allided with a bridge pillar in San Francisco Bay in dense fog. The transcription recollects the entire bridge team communication recorded over more than five hours (National Transportation Safety Board 2009).

The analysis focuses on the utterances by the pilot and the shipmaster as they constitute the majority of the total communication.

3.1. Definition of independent communication segments

The transcript of the published bridge team communication on board MV *Cosco Busan* stretches over a total time of 18,817s and contains 16,279 words. This leads to a ratio of 0.87 words/s. Although the utterances' length itself is not stated, the time difference between the beginning of an utterance and that of the subsequent one can be calculated. In the following, this time difference will be referred to as the utterance duration, abbreviated as UD. It spreads from 0 to 265 s and its arithmetic mean μ_{UD} is 8.3, i.e. on average, every 8.3 s a new utterance either by the same or by another speaker begins. The shortest utterance duration of zero seconds actually means that two or more people speak simultaneously. On average, each utterance is composed of 7.8 words, spreading from 1 to 157 words.

For a more detailed analysis of the communication flow before, during and after the allision, the data have to be divided into smaller communication segments. Given the limited vocabulary used in the very specific domain on board a ship, a segmentation

based on stylistic indices like vocabulary growth or vocabulary diversity (Labbé et al. 2004) or other methods based on type-token relations (Müller 2002) are not considered suitable because they are based on variations in language usage. In a setting where a high lexical variety (i.e. a lexically 'richer' way of communicating) may rather be considered an obstacle for an efficient information exchange, a segmentation on the basis of the duration of the different utterances is more appropriate. The following example shows that the type of communication and the wording does not change at all, the rudder commands are only interrupted by a period of silence.

Example 1: Communication at 08:15 LT (segments 21 and 22)	
CREW	Course 2/8/8 (segment 21)
PILOT	Steady (segment 21)
CREW	Steady (segment 21)
PILOT	2/8/5 (segment 22)
CREW	2/8/5 (segment 22)
CREW	Would you like something to eat? (segment 22)
PILOT	Um no do you have just some rice? (segment 22)

Assuming that UD presents an almost Gaussian distribution, 99.9% of all communication segments have a duration of less than $\mu_{UD} + 3*\sigma_{UD}$. As a working hypothesis, it is therefore assumed that a new communication segment starts if the time from the beginning of an utterance to the beginning of the next utterance is longer than this value. In the case of the *Cosco Busan* transcript, a communication segment is consequently defined as independent if its utterance duration is longer than 60.97s. Figure 2 displays the different communication segments over the time axis.

By applying this method, the bridge team communication recorded on MV *Cosco Busan* can be split into forty eight discrete segments. These have been validated by a qualitative analysis of their respective content. They all start with a clear indication of a new topic being introduced: a question is asked, a command is issued, an intention is expressed, communication changes from direct to radio communication or a remark is made following an action.

3.2. Non-weighted analysis

The words per time ratio presented above suggests a fluent and continuous communication situation. Nevertheless, it is quite unevenly distributed among the different speakers and segments as can be seen in Fig. 3 (Share of bridge team communication). The word total has been divided by ten to maintain the scale on the ordinate.

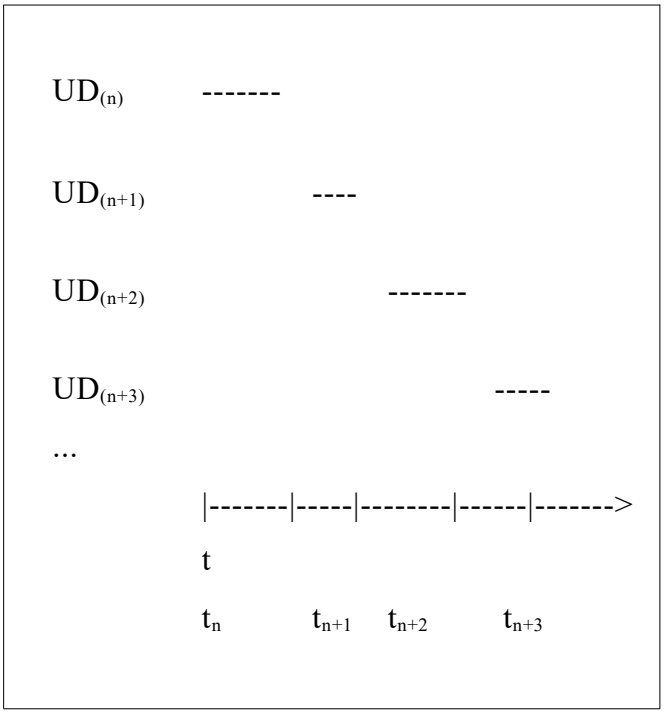

Fig 2: Communication segments over time axis

Over the whole sample, the shipmaster's word count represents 22.2% and the pilot's 27.2% of the entire communication, with both the shipmaster's and the pilot's utterances spreading from zero to one hundred percent of the segment analysed. The arithmetic mean values are 20.8 and 25.2%, respectively. After segment 36, the pilot disappears from the bridge altogether due to a sudden sickness following the impact, so there is no communication between the shipmaster and the pilot any longer. If the analysis is therefore limited to segments one to 36, the arithmetic mean of the shipmaster's utterances amounts to 17%, and the pilot's value rises to 33.6%. These general figures show that the pilot spoke nearly twice as much as the shipmaster.

This may be a normal distribution of communication streams between any pilot and shipmaster. However, in the analysis of different communication segments it becomes

quite obvious that a general analysis over the whole recorded time does not show that an exchange of information took place between the two bridge team members in the different situations before, during and after the allision. The interactions rather seem to occur in an asynchronous and unbalanced manner as only in segment twenty-seven the utterances by both speakers are equally distributed.

Although a mere word count is not suitable to evaluate any communication structure, it can serve as a first approximation towards a more comprehensive, weighted analysis.

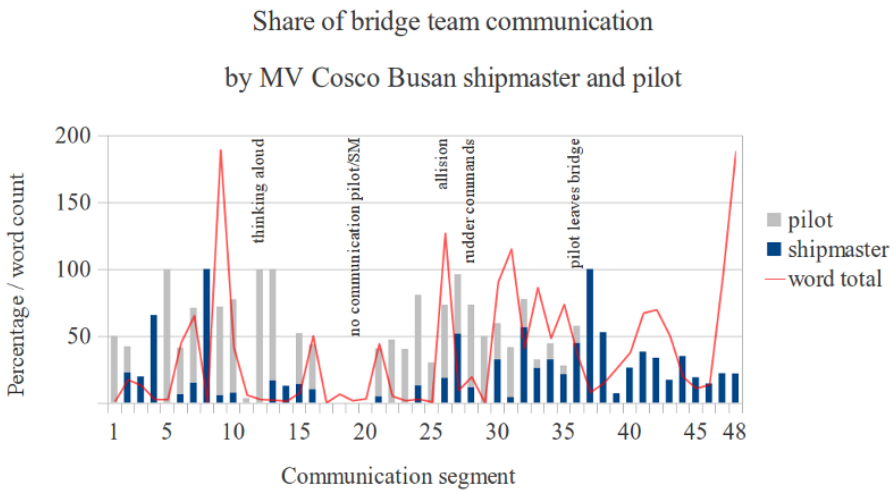

Fig 3: Share of bridge team communication

3.3. Weighted analysis

In a weighted analysis, some words are deemed to be more important functionally or lexically than others. As early as in 1891, Henry Sweet identified the distinction between the information content of different word classes:

> 'In a sentence such as *The earth is round*, we have no difficulty in recognizing earth and round as ultimate independent sense units .... Such words as *the* and *is*, on the other hand, though independent in form, are not independent in meaning: *the* and *is* by themselves do not convey any ideas, as *earth* and *round* do. We call such word as *the* and *is* form-words, because they are words in form only. When a form-word is entirely devoid of meaning, we may call it an empty word, as opposed to full words such as *earth* and *round*.' (Sweet 1891)

This has lead to the lexical density concept as one way of extracting the information

density of any given text. The lexical density (LD) is a percentage of words conveying information as opposed to those only fulfilling a grammatical function, the so-called functional words (also referred to as minor, empty, form, structural or grammatical words) or, as Halliday puts it, *the proportion of lexical items (content words) to the total discourse* (Halliday 1987).

Definition 1: lexical density $$LD = \frac{content\ words}{total\ word\ count} * 100$$

In this research, content words include nouns (e.g. ship and vessels) and proper nouns (e.g. *Cosco Busan*), adjectives (e.g. *big, bigger and biggest*) ordinal and cardinal numbers (e.g. *one, two, first* and *second*), adverbs (e.g. *slowly*) and full verbs (e.g. to *proceed* and to *depart*). Function words include all other word classes (e.g. pronouns (*he* and *them*), prepositions (*on*, *in* and *into*), determiners (*more*, *each* and *neither*), conjunctions (*and*, *nor* and *but*), modal and auxiliary verbs (*shall*, *may* and *need*).

Lexical density is a valuable tool to analyse the information content in any written text including transcribed conversations. Words are weighted according to their grammatical class but not to their content or appropriateness in a given situation. It is therefore not suitable to analyse any grade of nautical idiomaticity, i.e. this method alone is not able to show if the speakers use an appropriate, nautical vocabulary or not. The functional scoring of words needs to be complemented by a lexical weighting.

In order to achieve a lexical weighting of words, key words have to be identified which are scored higher than 'ordinary' words. In linguistics, a key word is defined statistically as a word which occurs more often in a given setting than we would expect to occur by chance alone. In bridge team communication, speech patterns have long been standardised by the Standard Marine Communication Phrases in an effort to implement a sort of artificial idiomaticity, or, as Schriever puts it: 'The SMCP is arguably the most constructive and widespread attempt to further understanding between international members in the maritime world. It is kept simple and easy to use.' (Schriever 2008, p. 39).

The SMCP are intended to be used as full sentences and pre-defined word patterns in English language to minimise any ambiguities due to an inconsistent communication. For this research, they have been used as a key word lexicon, though. All content words

of the SMCP have been isolated and used as key words which have a higher priority for the bridge team communication than any other words. Translations of the SMCP are also available for other major languages. Even if these are neither official nor to be used for communication purposes, they can still be applied as a lexicon in order to identify key words.

As a result, for assessing the information content of utterances three differently weighted word classes are counted: all words, content words and SMCP key words. These are counted for each speaker and time segment and divided by the total word count of all speakers for the given time segment. The resulting value is defined as the Individual Communication Index (ICI):

$$\text{Definition 2:} \quad ICI = \frac{n_s + n_{sc} + n_{sk}}{n}$$

where:

n is the total word count of all speakers for a given time segment.

n_s is the number of a given speaker's word count in a given time segment,

n_{sc} is the number of a given speaker's content words in a given time segment, and

n_{sk} is the number of a given speaker's SMCP key words in a given time segment.

For any given time segment, n is the total set of words used, n_s is a subset of n, n_{sc} a subset of n_s and n_{sk} a subset of n_{sc}. A different weighting is achieved by counting n and n_s once (score=1) while n_{sc} are counted twice (score=2) and n_{sk} are counted three times (score=3).

Figure 4 (Weighting of different word classes) displays the different word categories as sets and subsets.

The following two excerpts from the Cosco Busan VDR transcript show the different word classes with their respective weighting. Content words are highlighted in italics, SMCP key words are underlined.

Example 2: Communication at 07:19 LT (segment 9)	
PILOT	Does this *get AIS*? *AIS*?
CREW	Can you *see ship man[o]euvering information*?
CREW	I *see* it. I *see* it.
PILOT	No no no *okay*.
CREW	Well I didn't *[2].
PILOT	Oh *okay*. Some you can *see* that.
CREW	You can *see* the *signal*?
PILOT	It's been *selected*.
CREW	It's been *selected*.

The ICI for this situation would be calculated as follows:

$$ICI\ pilot = \frac{n_s + n_{sc} + n_{sk}}{n} = \frac{20+7+3}{50} = 0.6$$

$$ICI\ crew = \frac{n_s + n_{sc} + n_{sk}}{n} = \frac{30+10+10}{50} = 1.00$$

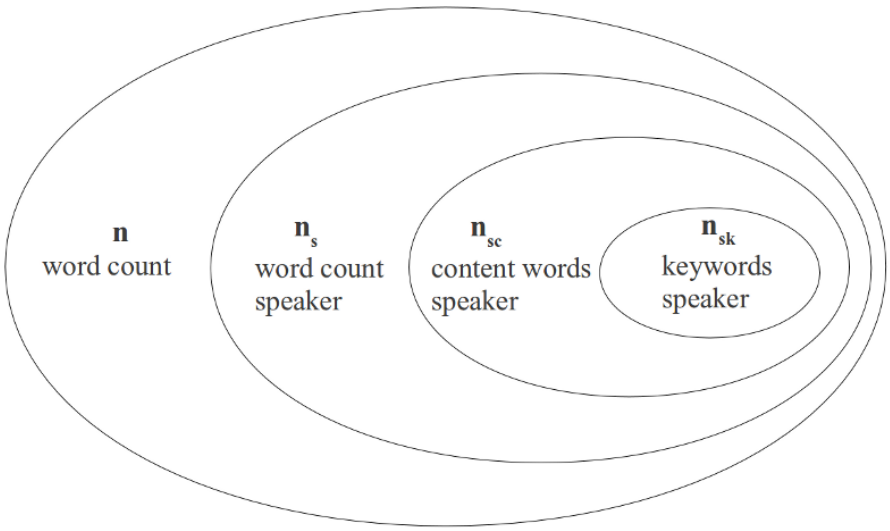

Fig 4: Weighting of different word classes

2 Unintelligible words form part of the 'all words' class and are weighted with the value 1

Example 3: VHF communication at 08:27 LT (segment 26)	
PILOT-VHF	*Traf[f]ic romeo* did you *call*
VTS	*Unit romeo traffic* uh *AIS shows* you on *two three five heading*
VTS	What are your *intentions over*?
PILOT-VHF	Well I'm *comin'* around I'm *steering two eighty*

In this case, the ICI values are:

$$ICI\ pilot = \frac{n_s + n_{sc} + n_{sk}}{n} = \frac{15+8+5}{32} = 0.86$$

$$ICI\ VTS = \frac{n_s + n_{sc} + n_{sk}}{n} = \frac{17+11+7}{32} = 1.09$$

These two examples show that the information content varies substantially among the different speakers and situations, ranging from 0.58 to 1.47. Fig. 5 (ICI for Shipmaster and Pilot) displays the communication indices in a bar chart over the communication segments so that the information content of the different speakers analysed can be compared visually.

It clearly points out the substantial differences in the pilot's ICI values and those of the shipmaster. However, at this point, it is still not clear which information density may be considered appropriate for an effective communication in a given situation on board a ship. The similarity of this diagram with Fig. 3 (Share of bridge team communication) rather suggests that the information content remains constant for a given speaker. In order to analyse if the information density in a bridge team member's utterances in fact remains unaltered throughout the different situations, the ICI variation over a speakers' mean ICI value μ ICI has been calculated for the various time segments:

Definition (3): $ICI\ variation = ICI_{CS} - \mu_{ICI}$

By calculating changes over a speaker's own average, his or her individual, idiosyncratic way of speaking is taken into account.

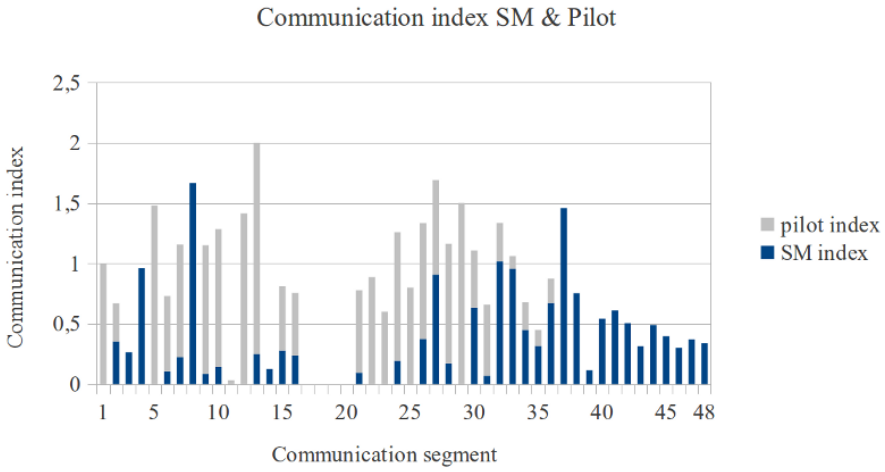

Fig 5: Communication Index for Shipmaster and Pilot

The bar chart in Fig. 6 (Communication index change over mean average) displays the variation of the individual ICI values by the shipmaster and the pilot over their respective arithmetic mean. The total height of the bars also shows the cumulated ICI variation, i.e. the change of the whole bridge team's information density.

3.4. Application of weighted analysis

Figure 7 (ICI values for exemplary situations) summarises some exemplary situations and links the individual and cumulated ICI variations to the situation taking place on board MV *Cosco Busan*.

Segment	Starting time	Situation	ICI SM (change)	ICI pilot (change)	Cumulated ICI change
...					
11	07:30:23	Assessment of situation: radar picture, external VHF communication, SM not speaking	0.00 (-0.35)	0.03 (-0.41)	-0.76
12	07:33:52	Pilot speaking to himself: assessment of navigational situation	0.00 (-0.35)	1.41 (+0.97)	+0.62
13	07:35:53	Questions by Pilot, SM answering → clear improvement, communication unbalanced	0.25 (-0.10)	1.75 (+1.31)	+1.21
14	07:37:15	Communication between SM &	0.13	0.00	-0.67

		crew → no communication between SM & Pilot	(-0.23)	(-0.44)	
15	07:39:01	Communication between chief officer & SM, crew & Pilot, VTS information → no clear communication streams, information not shared by bridge team members	0.28 (-0.07)	0.53 (+0.09)	+0.02
16	07:41:28	Communication between Pilot & crew, Pilot & VTS and other ship, SM & chief officer and crew → no communication between SM & Pilot	0.24 (-0.11)	0.52 (+0.08)	-0.03
17 to 20	07:55:04	Only communication by crew & chief officer and 2nd officer → no communication between SM & Pilot	0.00 (-0.35)	0.00 (-0.44)	-0.79
21 to 25	08:04:53	Communication by crew, SM, Pilot, rudder commands by SM, assessment of situation by bridge team members	0.06 (-0.29)	0.81 (+0.37)	+0.31
26 allision	08:24:49	Rudder commands with closed-loop communication between Pilot & crew, some interjections by SM, after allision SM more active	0.37 (+0.02)	0.96 (+0.52)	+0.54
27	08:40:30	Assessment of situation by SM & Pilot	0.91 (+0.56)	0.78 (+0,34)	+0.90
28	08:44:45	Rudder commands with closed-loop communication between Pilot & crew, interjections by SM	0.17 (-0.18)	0.99 (+0.55)	+0.37
29	08:49:06	Rudder commands with closed-loop communication between Pilot & crew	0.00 (-0.35)	1.50 (+1.06)	+0.71
30	08:50:35	Assessment of situation, Pilot disappears	0.63 (+0.28)	0.47 (+0.03)	+0.32
...					

Fig 7: ICI values for exemplary situation

It shows that both the individual and the cumulated indices vary substantially. The lowest value calculated for the shipmaster is 0.00 which equates to a change over his own arithmetic mean of −0.35. His highest value is 0.91, a change of +0.56 over his mean value. For the pilot, the values are 0.00 (−0.41) and 1.75 (+1.31), respectively.

It is of course quite obvious that the lowest cumulated values can be observed in situations where the shipmaster and pilot do not communicate with each other. Cumulated values oscillating around 0 occur where the speakers do not share information with the other bridge team members, e.g. in the following excerpt:

Example 4: Communication at 07:39:01 LT (segment 15)

CAPT-UHF	Chief officer it's all right to * anchor?
CHOFF-UHF	**The anchor is raised. *Deck status
CAPT-UHF	Okay
PILOT	Yeah it will go off all the time until he gets past us
CREW-UHF	Calling***
CREW	–*
PILOT	No I like it in the center. it would really #[3] me up if you did that

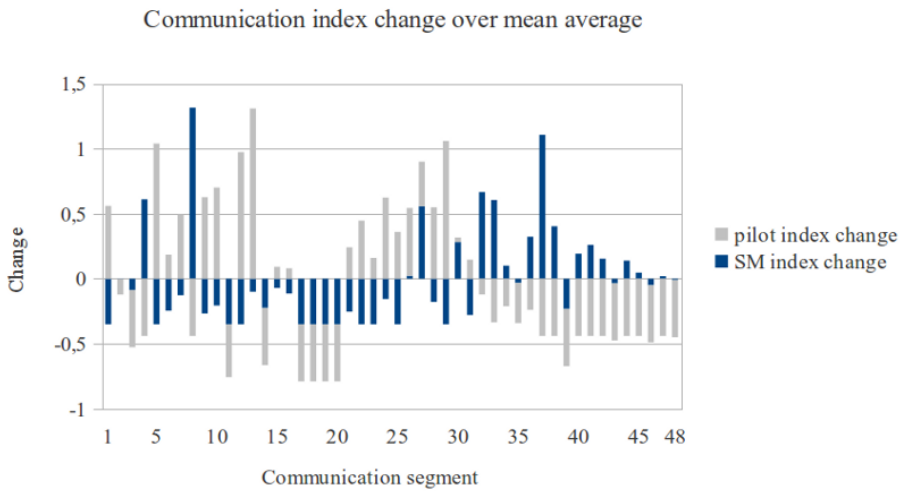

Fig 6: Communication index change over mean average

3 Expletive word

The time any reader needs to understand this situation replicates the situation on board MV *Cosco Busan* where bridge team members would have also required some time and effort to assess the situation. Had they shared their individual pieces of information with all the other members, the situational awareness of the group would have increased considerably. Instead, they were left to assess the overheard communication streams themselves, meaning that their individual language skills and navigational experience also led to situational interpretations which varied widely. Based on the assumption that the bridge team members were not able to understand the high percentage of unintelligible words any better than the transcription team, it is likely that communication was hindered further. Attorney John Meadows who represented the pilot in court summarised the situation as follows: 'While some information was exchanged, perhaps it could be said it wasn't a full transfer of information' (San Francisco Chronicle, 14 November 2007). This is clearly an understatement as the next conversational fragment suggests:

Example 5: Communication at 07:41:28 LT (segment 16)	
CREW	Other vessel*
PILOT	Pardon?
CREW	Other vessel*
PILOT	Oh yeah I think that as soon as he goes by yeah

The highest cumulated ICI values occur in segments 13 (+1.21), 27 (+0.90) and 29 (+0.71). In these situations unambiguous questions are answered (segment 13), the situation following the impact are assessed (segment 27) and rudder commands are given and answered in a closed-loop communication (segment 29). This last example clearly shows the difference with respect to example 4:

Example 6: Communication at 07:35:53 LT (segment 13)	
PILOT	And uh which direction do you go from the pilot station?
PILOT	North? south?
CAPT	North
PILOT	North channel? you take the north channel?
CAPT	Yeah yeah yeah

Here, a concise and unambiguous information exchange takes place which is reflected by the highest cumulated ICI change in the whole transcript (+1.21).

4. Conclusions

The main purpose of bridge team communication is the exchange of information which helps the navigators increase their situational awareness in order to ensure a safe navigation. A common understanding of the situation on board and the vessel's vicinity to navigational dangers by all team members is of crucial importance for manoeuvring any ship in a safe manner. Hence, *effective* communication plays a vital role for a safe ship operation. In an environment where verbal language is mainly used to share information, the information density can be directly related to the effectiveness of a given communication. Surprisingly though, no method has yet been defined to analyse communication structures suitable to differentiate communication from miscommunication. This article has introduced a quantitative methodology to calculate and weight utterances for evaluating its information content. Such a method can be applied regardless of the language used by the speakers because of its quantitative nature and its emphasis on grammatical and lexical weighting.

The limited size of the sample and the high percentage of unintelligible words (1,242 in total) is clearly not sufficient to present statistically sound findings. Nevertheless, even this small sample leads to the following conclusion:

The quality of the information content in bridge team communication is measurable.

Asynchronous or unbalanced communication streams can be detected using a simple word count. By weighting different words functionally and lexically the information quality can be extracted for different speakers, time segments and situations. An evaluation is possible on an individual as well as on a team level. Alterations in communication patterns due to modified external (situational) or internal (personal) conditions can be identified. This leads to the following hypothesis:

The higher the cumulated ICI, the better the bridge team communication.

This preliminary study suggests that *effective communication* occurs where the cumulated ICI is above average whereas *ineffective communication* takes place in situations with a cumulated ICI far below average. A working hypothesis is that the

higher the cumulated ICI, the more effective the bridge team communication and, vice versa, the lower the cumulated ICI, the more likely is an ineffective communication. However, as the reference points are the speakers' own average values, these will have to be considered for the different navigational situations (e.g. anchoring, rudder commands, search and rescue scenarios) as well as taking into account the different speakers' background (mother tongue, experience on board ship, etc.).

There is still ample scope for validating the presented methodology in the framework of a large, specific text corpus for the nautical environment with an aim to eventually improve on-board communication and thus to reduce human errors based on ineffective communication.

References

Bowers CA, Jentsch F, Salas E, Braun CC (1998) Analyzing communication sequences for team training needs assessment. Human Factors 40(4):672–672

Chauvin C (2011) Human factors and maritime safety. J Navig 64:625–632

Chin K-S, Pun K-F, Ho ASK, Lau H (2002) A measurement-communication-recognition framework of corporate culture change: an empirical study. Human Factors and Ergonomics in Manufacturing 12:365–382

Chow, R., Christoffersen, K., & Woods, D. D. (2000) A model of communication in support of distributed anomaly response and replanning. Human Factors and Ergonomics Society Annual Meeting Proceedings 44.1: 34.

Cole, C. (Ed) (1994) A Yardstick of English language competence for ship's officers. Echo 29/December German Association for Maritime English, p. 11.

Cole CW, Trenkner P (2009) The Yardstick for Maritime English STCW assessment purposes. IAMU Journal 6(1):13–28

Cole, C. W., Trenkner, p. (2012) Whither Maritime English?—2012. In: Proceedings of the 24nd International Maritime English Conference, pp. 3–18

Cooke NJ, Salas E, Cannon-Bowers JA, Stout RJ (2000) Measuring team knowledge. Human Factors 42(1):151–173

Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Human Factors 37:32–64

Froholdt, L.L. (2010) Getting closer to context: a case study of communication between ship and shore in an emergency situation. Text & Talk 30.4:385+.

Gonzalez C (2005) Task workload and cognitive abilities in dynamic decision making. Human Factors 47(1):92–101

Grech, M.R. (2005) Human error in maritime operations: assessment of situation awareness, fatigue, workload and stress. Doctoral thesis, The University of Queensland

Halliday, M.A.K. (1987) Spoken and written modes of meaning. In: Horowitz, R. & Samuels, S.J. (eds) Comprehending Oral and Written Language. Academic, San Diego

Hobbs A, Williamson A (2003) Associations between errors and contributing factors in aircraft maintenance. Human Factors 45(2):186–201

Hutchins, E. (1995) Cognition in the wild. MIT Press, Cambridge.

International Maritime Organization (1978). Standards of training, watchkeeping for seafarers, 1978 as Amended, London.

International Maritime Organization (2001) Resolution A.918(22): IMO Standard Marine Communication Phrases, London.

International Maritime Organization (2004). SOLAS consolidated edition 2004, Chapter V, Regulation 14/3, London.

International Maritime Organization (2010). Conference of Parties to the STCW Convention 1978, Conference Resolution I, Regulation I/14/.7, London.

Jentsch F, Barnett J, Bowers CA, Salas E (1999) Who is flying this plane anyway? What mishaps tell us about crew member role assignment and aircrew situation awareness. Human Factors 41(1):1–14

Khawaja MA, Chen F, Marcus N (2012) Analysis of collaborative communication for linguistic cues of cognitive load. Human Factors 1:2012

Kim S, Park J, Kim YJ (2011) Some insights about the characteristics of

communications observed from the off-normal conditions of nuclear power plants. Human Factors and Ergonomics in Manufacturing and Service Industries 21(4):361–378

Koiso T, Nishida S (1999) Communication support system for operators in emergency of large-scale plant. Human Factors in Manufacturing 9(4):357–366

Labbé C, Labbé D, Hubert P (2004) Automatic segmentation of texts and corpora. Journal of Quantitative Linguistics 11(3):2004

Levin (2006): Levin, S., Hemphill, D.J., Jones, I., et al. (2006) Tracking workload in the emergency department. Human Factors 48(3):526–39.

The Marcom Project (1999) The impact of multicultural and multilingual crews on maritime communication. Contract No WA-96-AM-1181. A Transport RTD Programme DG VII.

Marine Accident Investigation Branch (2005) Report on the investigation of the collision between Lykes Voyager and Washington Senator, London.

Marine Accident Investigation Branch (2007) Annual Report 2007, London.

Marine Accident Investigation Branch (2008) Annual Report 2008, London.

Marine Accident Investigation Branch (2009) Annual Report 2009, London.

Marine Transportation Research Board (1981) Research needs to reduce maritime collisions, rammings, and groundings, Washington.

McCallum, M., Raby, M., Forsythe, A.M., Rothblum, A.M., Smith, M.W. (2000) Communications problems in marine casualties: Development and evaluation of investigation, reporting and analysis procedures. In: Proceedings of the Human Factors and Ergonomics Society, Annual Meeting, 4.

McCrae C (2009) Human factors at sea: common patterns of error in groundings and collisions. Maritime Policy and Management 36(1):21–38

Mishra D, Mishra A (2009) Effective communication, collaboration, and coordination in extreme programming: human-centric perspective in a small organization. Human Factors and Ergonomics in Manufacturing 19(5):438–456

Müller D (2002) Computing the type token relation from the a priori distribution of types. Journal of Quantitative Linguistics 9(3):193–214

National Transportation Safety Board (2009) Accident report NTSB/MAR-09/01PB2009-916401, Washington.

Paramore, B., Dayton, R.B., Porricelli, J.D., & Willis, R.M. (1979) Human and physical factors affecting collisions, rammings, and groundings on western rivers and gulf intercoastal waterways, vol. 1 (CG-D-30-78) Groton.

Pyne, R. & Koester, T. (2005) Methods and means for analysis of crew communication in the maritime domain. The Archives of Transport, vol. XVII, No. 3–4.

Quinn PT, Scott SM (1982) The human element in shipping casualties: Analysis of human factors in casualties. London: Tavistock Institute of Human Relations.

Rodrigues de Carvalho P, Benchekroun TH, Gomes JO (2012) Analysis of information exchange activities to actualize and validate situation awareness during shift changeovers in nuclear power plants. Human Factors and Ergonomics in Manufacturing and Service Industries 22(2):130–144

Salas E, Wilson C, Burke A, Shawn W, Dennis C (2006) Does crew resource management training work? An update, an extension, and some critical needs. Human Factors 48(2):392–412

Salas E, Fowlkes JE, Stout RJ, Milanovich D, Prince C (1999) Does CRM training improve teamwork skills in the cockpit? Two evaluation studies. Human Factors 41(2):326–343

San Francisco Cronicle, 14 November 2007. Pilot says Cosco Busan's Captain directed vessel into bridge, San Francisco.

Schriever, U. (2008) Maritime communication in an international and intercultural discourse. Doctoral thesis, The University of Tasmania.

Trenkner, P., Nielsen, D. (1998). A Method for the Acquisition of Statistical Data on Maritime Communication Deficiencies. Project under the Germany/Hong Kong Joint Research Scheme 1997/98. Hong Kong: Warnemuende

Trenkner, P. (2002). The IMO-Standard Marine Communication Phrases and the Requirements of the STCW Convention 1978/95, Proceedings of the International ITUMF-JICA Seminar on Maritime English (pp. 37–57). Istanbul Technical University, Turkey.

Trenkner, P., Cole, C. (2010). Raising the Maritime English bar: The STCW Manila Amendments and their impact on Maritime English. In: Proceedings of the 22nd International Maritime English Conference, pp. 3–16.

Sweet, H (1891) A New English grammar: logical and historical. Clarendon Press, London.

UK Department of Transport (1991) The human element in shipping casualties. UK Department of Transport, London.

Webster J, Cao CGL (2006) Lowering communication barriers in operating room technology. Human Factors 48(4):747–758

Lingua Franca and its Grammar Footprint: Introducing an Index for Quantifying Grammatical Diversity in Written and Spoken Language[1]

Peter John

Faculty of Maritime Studies, Jade University of Applied Sciences, Germany

Australian Maritime College, University of Tasmania, Australia

Benjamin Brooks

Australian Maritime College, University of Tasmania, Australia

Abstract

The paper introduces a method to quantify grammar diversity based on part-of-speech tagging (POS) and presents a POS diversity index which can be used for analysing written and spoken communication independently of a text's length. For this purpose, a series of transcripts of the radio programme *Lingua Franca*, broadcast by the Australian Broadcast Corporation, are assessed. The radio presenter's grammatical footprint is compared with the POS index values of the invited guests, and the hypothesis that a significant difference in the grammar structures used by the different speakers is tested for validity using a standard ANOVA calculation. Significantly different POS diversity is found in a comparison between groups of speakers. The index provides for comparative studies of grammar structures used by individual speakers or authors.

1. Introduction

A range of linguistic indices have been developed to carry out research on written texts. Indices aim at quantifying observations in language in order to compare the number of occurrences in texts produced by different authors or by the same author. They are used

1 This is an Accepted Manuscript of an article published by Taylor & Francis in the Journal of Quantitative Linguistics on 17 December 2013, available online: http://www.tandfonline.com/10.1080/09296174.2013.856130.

for identifying literary genres and for analysing grammatical and syntactical features. They also assist in explaining and interpreting texts, e.g. in the exegesis of religious and secular texts. In the field of language research, quantified linguistic observations can be correlated with internal or external variables. Internal variables refer to other observations prevalent in the studied document whilst external variables include information collected outside the studied text, such as the author's gender, age, status and mother tongue.

This paper introduces the specific POS diversity index (spdi) which provides for a simple and robust way to quantify the number of different grammar classes used on the basis of part-of-speech tagging (POS). The index is applied to the transcripts of 10 radio programmes, and a comparison of the POS diversity as used by different speakers is carried out.

2. Literature review

Grammar complexity has been studied by a multitude of researchers, most notably in the field of first language acquisition by children, foreign language learning and in the area of speech impairment. Nevertheless, the number of indices and procedures used for studying grammar diversity is rather limited.

Grammatical complexity has been addressed by using a word count system by Brown (1973) who proposed the measurement of the mean length of utterances in morphemes (MLU). This "excellent simple index of grammatical development" counts the number of morphemes in spontaneous utterances as a measure of grammatical complexity. It is a commonly used metrics in linguistic and psycholinguistic studies to calculate the degree of grammar complexity. However, by simply counting morphemes (or words) no changes are detected when words or grammar structures are repeated in an utterance, which may cause inaccurate suppositions on grammar complexity in utterances with repeated elements like, for example, "I don't know, I really don't know".

Other metrics for estimating grammar complexity carry out a syntactic analysis by counting the number of embedded or subordinate clauses. These include:

- The Mean Clauses per Utterances (MCU), by Kemper et al. (1989) and Cheung and Kemper (1992).

- The Developmental Sentence Score (DSS) by Lee and Koenigsknecht (1974) which identifies eight categories of grammatical forms and assigns points to each sentence.
- The Developmental Level metrics (DLevel) which was developed by Rosenberg and Abbeduto (1987) and modified by Cheung and Kemper (1992). It distinguishes between eight levels of complexity looking at the type and number of subordinate phrases and clauses.
- The Index of Productive Syntax (IPSyn) by Scarborough (1990) which counts phrases and sentence structures.
- The Propositional Density (PropD) which is used as a measurement of content.
- Procedures for scoring propositional content by Kintsch and Keenan (1973) and by Turner and Green (1977).
- Analyses of syntactic complexity by Blake and Quartaro (1990) by summing up identified clause units for a given utterance.

Compared with the word-based MLU, all indices based on syntactic analysis require a thorough and time-consuming identification of the sentence structure to be coded.

Another tool for analysing grammar structure is the Language Assessment, Remediation and Screening Procedure (LARSP) introduced by Crystal et al. (1976; Crystal, 1982). It aims at a "developmental description of children's language which is based on the grammatical framework of an adult reference grammar" (Klee, 1985). It has been developed for a series of languages and is commonly used in language assessment.

3. Methodology

In this paper, grammar diversity is studied at three different levels. Firstly, the ratio of grammar classes over utterance length is calculated. Secondly, an expected number of grammar classes is identified for different utterance lengths by means of a regression analysis and thirdly, the individual speakers' actual grammar diversity is compared with the expected values. Figure 1 details the different analysis levels:

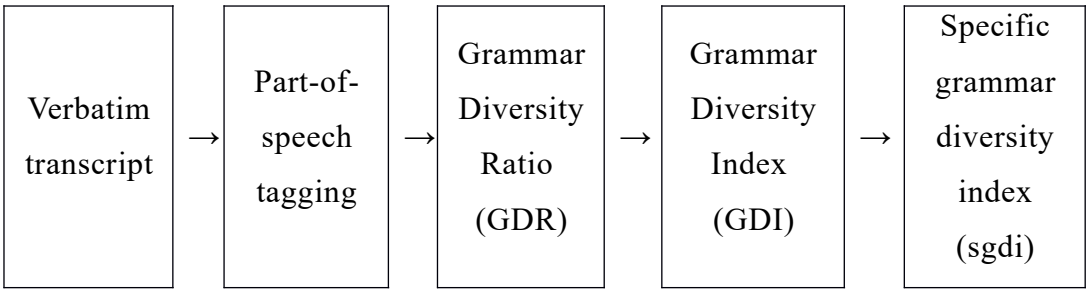

figure 1: Analysis levels of POS diversity

3.1 Data Sampling and Definition of Grammar Diversity

The verbal communication analysed in this paper is based on verbatim transcripts of 10 editions of the radio programme Lingua Franca, broadcast by the Australian Broadcast Corporation (ABC) from 18 February to 21 April 2012. According to its Internet homepage, "[t]he program features experts who analyse a single topic of interest to users and lovers of language. Examples of the sort of linguistic territory they traverse are: bi-lingual education, ebonics, the language of pornography, and the political use of words" (ABC 2013). Presenter Maria Zijlstra invites guests to speak about a variety of topics related to communication. Each programme has an approximate duration of 14 minutes during which one or two guests are interviewed.

Analysing a transcribed radio programme completely eliminates the impact of the observation on participants, thus warranting an optimal verbal language production without any biases introduced by a present observer. In the *Lingua Franca* programme, the presenter interacts with her guests in a spontaneous and natural way as she asks questions and encourages her guests to continue speaking, go into more detail, etc. Out of the 76 questions the presenter asks in the 10 editions studied, the following questions can serve as examples for the distinctly verbal character of the interviews: "They are all things that he wrote about, right?" (18 February), "The indigenous languages, you mean?" (25 February), "Can you elaborate, please?" (17 March), "It seems almost absurd, right?" (24 March).

The 10 transcripts of the *Lingua Franca* radio programme offer 10 observations for the presenter and 12 observations for the invited guests as eight editions include one guest only and two editions two guests. The transcripts include a total of 20,547 words out of which 4234 are produced by the presenter (M = 423, SD = 163) and 16,313 words by the

12 invited guests (M = 1,359, SD = 597). The transcripts are available on the programme's Internet pages. They do not contain any unintelligible words and are assumed to be a true reproduction of the broadcast radio interviews. Table 1 presents the word count for the 10 editions.

date	18/02	25/02	03/03	10/03	17/03	24/03	31/03	07/04	14/04	21/04
presenter	414	387	493	512	588	420	141	191	414	674
guest(s)	1712	1339	1927	1332	1451	1585	2059	1754/296	1249/96	1513

table 1: Word count presenter/guest(s)

In order to quantify the grammatical structure of the utterances, the different grammar classes were identified by the researchers with the robust C5 part-of-speech (POS) tagging method which was also used to tag the 100-million-word British National Corpus (Leech et al., 1994; Garside, 1996). The C5 tag set differentiates between 62 grammar classes (e.g. NN0 neutral nouns (*aircraft*, *data*), NN1 singular nouns (*table*, *goose*), NN2 plural nouns (*tables*, *geese*), NP0 proper nouns (*Paris*, *Peter*, *Pluto*); AJ0 unmarked adjectives (*good*, *old*), AJC comparative adjectives (*better*, *older*), AJS superlative adjectives (*best*, *oldest*)) (cf. Ucrel, 2013).

Each of the 970 utterances of the Lingua Franca transcript was analysed separately for the occurrence of the 62 grammar classes contained in the C5 tag set. In the case a grammar class appeared more than once, the double occurrence was not counted because a grammar class which appeared repeatedly was assumed not to contribute to a higher grammar diversity in the sense of a "state or quality of being different or varied" (Collins, 2006).

For each utterance a POS Diversity Ratio or PDR was computed as the number of the different grammar classes over the utterance length (total amount of words used in that utterance). n w stands for utterance length (word count of the utterance) and n GC for the number of different grammar classes used in an utterance. For the purpose of this paper, the term "utterance" is used synonymously for "sentence".

$$PDR = \frac{grammar\ class\ count}{utterance\ lengt\ h} = \frac{n_{GC}}{n_w} \quad (1)$$

Table 2 illustrates the way the utterance length and grammar classes are counted. 1[2]

The PDR adopts values of > 0 and ≤ 1 regardless of a text's length. A value near 0 occurs when a relatively long text only consists of one grammar class, e.g. an enumeration of numbers (1, 2, 3, 4, 5, ...). A PDR value of 1 is the result of a text in which no grammar class is used twice in each utterance and which lexical diversity is therefore very high.

example 1	Go	ahead.								n_{w1}=2
class	VVB	AV0								
GC	1	1								n_{GC1}=2
										GDR_1= 1.000
example 2	She	is	very	angry.						n_{w2}=4
class	PNP	VBZ	AV0	AJ0						
GC	1	1	1	1						n_{w2}=4
										GDR_2=1.000
example 3	So	I	still	have	a	long	way	to	go.	n_{w3}=9
class	AV0	PNP	AV0	VHB	AT0	AJ0	NN1	TO0	VVI	
GC	1	1	0	1	1	0	1	1	1	n_{GC3}=7
										GDR_3= 0.778
example 4	What	have	you	said	to	me?				n_{w4}=6
class	DTQ	VHB	PNP	VVN	PRP	PNP				
GC	1	1	1	1	1	0				n_{GC4}=5
										GDR_4=0.833
example 5	It	is	easy	and	it	is	astonishing.			n_{w5}=7
class	PNP	VBZ	AJ0	CJC	PNP	VBZ	AJ0			
GC	1	1	1	1	0	0	0			n_{GC5}=4
										GDR_5=0.571

table 2: Utterance length and grammar class (POS) count.

However, it does not take into account the number of words a text consists of. Example 1 and 2 of Table 2 both have a PDR value of 1 although Example 1 contains two words and Example 2 contains four words: PDR 1 = 2:2 = 1, PDR 2 = 4:4 = 1. Example 3 is

2 A list of grammar classes included in the C5 tag set is displayed on the Internet at: http://ucrel.lancs.ac.uk/claws5tags.html

comprised of nine words and eight different grammar classes which leads to a PDR 3 = 7:9 = 0.778. The values of Example 4 (six words, five grammar classes) and Example 5 (seven words, four grammar classes) are 0.833 and 0.571, respectively. In the Lingua Franca programme broadcast on 18 February 2012 the presenter produces 414 words with a mean PDR of 0.689, her guest's word count is 1712 with a mean PDR of 0.599. Table 3 displays the mean PDR values for the 10 radio programmes.

Some text-based ratios vary with the length of the texts they are based on as in the case of the type-token ratio (TTR) which decreases in line with the text length because the vocabulary size is not endless (cf. Herdan, 1960, 1966; Covington et al., 2010; Müller, 2002). A comparison of different text lengths is therefore difficult if not impossible. The POS diversity ratio presented in this paper is also subject to a decrease in the ratio of grammar classes over the utterance length. This is due to the fact that the sum of the different grammar classes can never exceed the total number of grammar classes identified. In the case of the C5 tag set the maximum grammar class amount equals 62.

date	18/02	25/02	03/03	10/03	17/03	24/03	31/03	07/04	14/04	21/04
presenter	0.68	0.53	0.46	0.65	0.70	0.78	0.51	0.53	0.79	0.76
guest(s)	0.59	0.63	0.68	0.63	0.64	0.63	0.61	0.67/ 0.73	0.72/ 0.78	0.76

table 3: Mean GDR values of presenter and guest(s)

It also does not differentiate between short and long utterances, as shown in the examples above. The utterances "I am hungry" and "I am definitely not willing to speak up" both present an identical grammar diversity ratio(PDR = 1) as each grammar class appears only once.

To determine if the PDR correlates with the number of words per utterance, Pearson's correlation coefficient was calculated for the 970 utterances of the transcript. The PDR values over utterance length lead to a strong negative correlation of −0.832, leading to the assumption that as a trend a longer utterance length would lead to a lower PDR value. The coefficient of determination R 2 = 0.692.

The next step was to group PDR values according to their text length and to calculate their means. This was carried out up to an utterance length of 25 words. Longer

utterances were not considered due to their sporadic occurrences. Table 4 summarizes the mean PDR values with their respective standard deviations (SD).

utterance length	mean GDR	SD	utterance length	mean GDR	SD
1	1.00	0.00	14	0.76	0.07
2	0.94	0.16	15	0.72	0.13
3	0.94	0.13	16	0.69	0.10
4	0.95	0.10	17	0.70	0.10
5	0.94	0.10	18	0.65	0.10
6	0.85	0.15	19	0.69	0.10
7	0.86	0.10	20	0.66	0.09
8	0.83	0.13	21	0.65	0.09
9	0.79	0.12	22	0.59	0.09
10	0.82	0.13	23	0.61	0.08
11	0.77	0.13	24	0.57	0.09
12	0.77	0.11	25	0.57	0.07
13	0.77	0.10			

table 4: Mean GDR for utterance lengths up to 25 words

As shown in Figure 2 a gently declining exponential regression can be observed in the relationship between the dependent variable “PDR” on the y axis and the independent variable “words per utterance” on the x axis.

In order to calculate the regression three assumptions have to be made. Firstly, the independent variable has to be known exactly for all elements of the sample. This is certainly the case when counting the number of words in each utterance. The second assumption refers to the distribution of the dependent variable around its mean value which should be normal. For an utterance with one word only, the PDR always adopts a value of one. This is displayed in Table 4 where the standard deviation for a one-word utterance is 0.000.

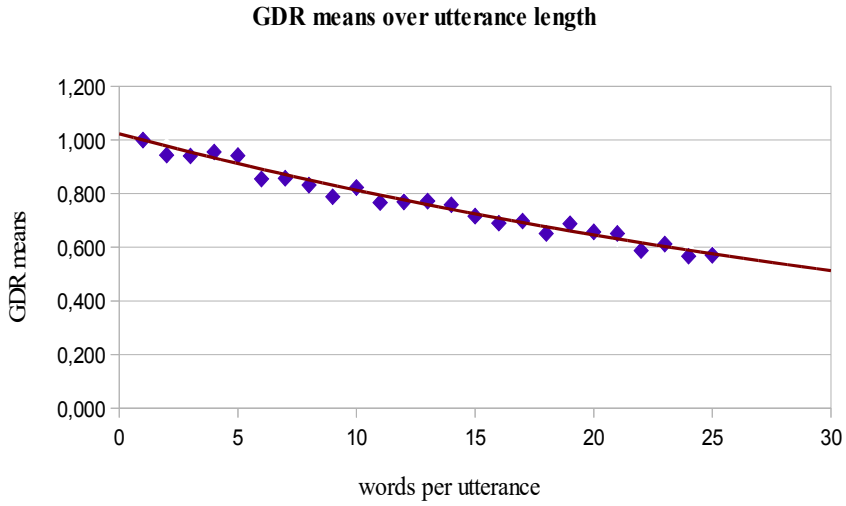

figure 2: PDR means over utterance length

For utterance lengths from two to 25 words, a Shapiro-Wilk normality test resulted in p = 0.973, a standard distribution can therefore be taken for granted. The third assumption relates to the standard deviation of the dependent variable, which should be identical for all points of the line. Based on the figures displayed in Table 4 the PDR's standard deviation for the different utterance lengths is assumed to be small enough to be accepted.

An utterance length of one word permits a PDR value of one only, because one word can obviously not have more or less than one grammar class. This means that the regression line needs to pass the point (1;1), thus reducing the regression calculation by one degree of freedom. Hence the non-linear combination of parameters was computed by Formulae (2) and (3):

$$PDR = \exp\left(-\alpha\left(n_w - 1\right)\right) \quad (2)$$

$$f\left(n_w\right) = \ln\left(PDR\right) = -\alpha\left(n_w - 1\right) \quad (3)$$

The regression of the line was calculated by applying the least squares method as an iterative procedure to find a value that best fits the data, leading to a coefficient α = 0.023 as the nearest approximation. Based on the exponential correlation of the dependent variable (PDR) with the independent variable (words per utterance) and by applying Formula (4), a typical PDR value can be calculated for utterances with a

determined number of words. This expected PDR value shall be defined as the POS Diversity Index or PDI.

$$PDI = \exp\left(-0.023\left(n_w - 1\right)\right) \quad (4)$$

Since it is quite natural that with increasing sentence size ever more categories are repeated, it is evident that the change of the variable y (=PDI) is negatively proportional to its topical value, hence $y' = -by$, yielding the differential equation whose solution is the function presented by the authors.

The exponential regression presents a very high coefficient of determination $R^2=0.972$ which indicates that the regression line fits the data.

The PDI expresses the expected amount of different grammar classes in an utterance. It is the nearest approximation to the values observed in Table 4 based on a least squares calculation of the population's mean.

Now it is also possible to compare the number of different grammar classes observed in a sample as opposed to the PDI as its expected value. By dividing the PDR by the PDI values of an utterance, a ratio between the expected and the observed number of grammar classes of an utterance is obtained. This ratio shall be defined as the Specific POS Diversity Index or spdi as expressed in Formula (5):

$$spdi = \frac{PDR}{PDI} = \frac{\frac{n_{GC}}{n_w}}{\exp -0.023\left(n_w - 1\right)} = \frac{n_{GC}}{n_w \exp\left(-0.023\left(n_w - 1\right)\right)} \quad (5)$$

The spdi adopts a value of 1 if the observed number of grammar classes equals the expected PDI. A value of 2 is obtained if the number of observed grammar classes is twice as high as the expected PDI, and a value of 0.5 is given if the number of observed grammar classes is half the expected PDI.

3.2 Data Analysis and Validation

A linguistic index is only applicable for comparing different speakers, authors, text types and so on, if it is able to make evident statistically significant differences between the speakers, authors and text types studied. It is understood that both the presenters and the guests' grammatical diversity values need to be distributed in a standardized, bell-

shaped curve if their variances are to be compared for any divergence. The 10 observations for the presenter lead to a mean spdi of 1.01 with a standard deviation (SD) of 0.22. With a 95% confidence interval, the lower bound is 0.98 and the upper bound 1.04. The values for the 12 observations of the guests' utterances result in a mean spdi of 1.07, and a SD of 0.22. The lower and upper bounds are 0.98 and 1.04, respectively.

In order to test if the index adopts different values for the individual speakers, two statistical analyses are performed. The first one analyses the variance of the presenter's spdi values with that of the invited guests as a group. The second analysis compares the presenter's spdi variance individually with those of each of the invited guests.

For the first analysis, the null hypothesis states that no significant difference can be observed between groups, i.e. between the presenter's and the guests' spdi values, hence: H_0:spdi $\text{mean}_{\text{presenter}}$ = spdi $\text{mean}_{\text{guests}}$. Alternatively, a statistically significant difference in the spdi can be observed between the two groups: H_1:spdi $\text{mean}_{\text{presenter}} \neq$ spdi $\text{mean}_{\text{guests}}$.

Two statistical groups were examined by way of a one-way analysis of variance (ANOVA). The programme's presenter and the invited guests were used as fix factors (presenter and other speakers) with the spdi values as their dependent variables.

As a pre-requisite to carrying out an analysis of variance on the data, its homoscedasticity assumption has to be tested in order to make valid statistical inferences on population relationships. Levene's test for homogeneity of variances (Levene, 1960) was carried out for the 970 dependent variables. It resulted in $F_{1,969}$ = 0.05, p = 0.942 which warranted a sufficient homogeneity at a 95% confidence level and lead to the assumption that the variables' variances were sufficiently similar to carry out an analysis of their variance.

The ANOVA itself was performed with two fixed factors (presenter, guests) as a nominal, independent variable and the presenter's and guests' spdi values as dependent variable. It turned out to be highly significant with $F_{1,968}$ = 13.473 and p = 0.0002, so the null hypothesis can be rejected and the alternative hypothesis accepted. The presenter's spdi values differ significantly from the values observed in a group of guests. Figure 3 displays the mean spdi values as a box plot.

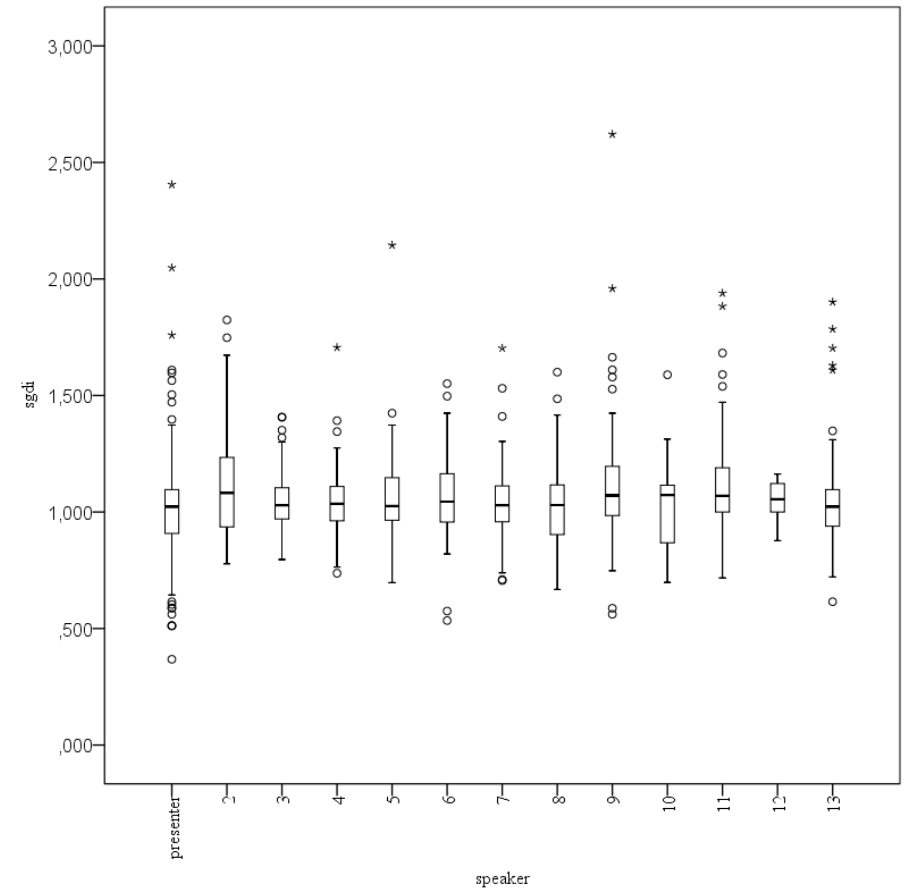

figure 3: Spdi values by presenter and guests

ANOVA conducted on their spdi values were below a 5% significance level (numbers 2, 9 and 11 in Figure 2), for six guests it was below 10% (numbers 3 to 5, and 13) and for another three guests it was 10% (numbers 8, 10 and 12). The box plot display in Figure 3 shows that despite the lacking homoscedasticity and a significance level of 10% of all participants, clear differences between the expected and the observed use of different grammar forms per sentence are visualized for the individual speakers.

As a cross-check a comparison of the presenter's spdi values in the different editions of the programme were compared. In this case the null hypothesis states that no significant difference can be observed between the presenter's spdi in the different radio programmes: H_0:spdi $\text{mean}_{\text{presenter 18th February}}$ = spdi $\text{mean}_{\text{presenter 25th February}}$ = ... = spdi $\text{mean}_{\text{presenter 21st April}}$. H_1:spdi $\text{mean}_{\text{presenter 18th February}}$ ≠ spdi $\text{mean}_{\text{presenter 25th February}}$ ≠ ... ≠ spdi $\text{mean}_{\text{presenter 21st April}}$. The ANOVA carried out lead to $F_{1,9} = 0.64$, $p = 0.74$. Levene's test for homogeneity of variances resulted in $F_{1,9} = 1.99$, $p = 0.05$ so the samples' variability has

to be considered heteroscedastic at a 0.05 level. This underlines the assumption that samples of utterances by the same person in the same condition (e.g. in this radio programme setting) do not vary significantly.

4. Discussion

This paper quantifies grammar diversity at three levels. The POS Diversity Ratio (PDR) is the number of grammar classes in a given utterance divided by the word count of the utterance. It has been demonstrated that the PDR does not grow in line with the text length. For comparing different texts an identical number of words needs to be chosen.

By displaying the PDR value over the word count of an utterance, an exponential regression can be calculated which leads to an expected PDR for a given utterance length counted in words. This POS Diversity Index (PDI) can be used for comparing the expected number of different grammar classes with the number actually observed. The PDI has been obtained on the basis of a population of 970 utterances including a total of 20,547 words.

The specific grammar diversity index (spdi) shows the relative deviation of the grammar diversity value in an observed utterance from the expected number given by the PDI for a specific utterance length. A spdi value of 1 is obtained when the observed and the expected values are identical, a spdi value of 2 means that the utterance contains twice as many grammar classes as expected, and a value of 0.5 shows that the utterance's grammar diversity is half of the expected value. The spdi index is independent from the utterance length and can be used to study the grammar usage for a given speaker or author in different situations, text types, etc. It can also be used to discriminate the grammar diversity by different speakers or authors.

Based on the example of the *Lingua Franca* radio programme it can be demonstrated that a between groups comparison can lead to an idiosyncratic spdi value for a given speaker. This grammar footprint can be compared with the values obtained for other speakers. It can also be used horizontally in research comparing different text types or discriminating spoken from written language, amongst other possibilities. The POS Diversity Index also provides for vertical or within groups studies, e.g. a person's speech patterns over a time axis, in different contexts, etc. Correlations with other ratios used in

linguistics such as the type-token ratio (TTR), mean length of utterances (MLU), individual communication index (ICI) and others can be studied.

By quantifying grammar diversity on the basis of part-of-speech (POS) tagging a simple and robust method has been presented which can be used along with other procedures using the same POS data for quantifying other linguistic features of written and spoken language (cf. John et al., 2013). It is more accurate than the MLU as it takes into account word repetitions, and less complex than syntactic procedures because it is based on a part-of-speech tagging method. It opens up a broad range of comparative linguistic research into grammar usage and provides unambiguous quantitative measures for comparative text analyses.

References

ABC, Australian Broadcast Corporation (2013). Lingua Franca – Program Podcast. Retrieved 13th March 2013, from http://www.abc.net.au/radionational/feed/2883984/podcast.xml

Blake, J., & Quartaro, G. (1990). Manual for recording, transcribing and analyzing preschool children's speech samples Report No: York University Department of Psychology. 189.

Brown, R. (1973). A First Language: The Early Stages Cambridge, MA: Harvard University Press.

Cheung, H., & Kemper, S. (1992). Competing complexity metrics and adults' production of complex sentences. Applied Psycholinguistics, 13, 53–76.

Collins, W., & Anderson, S. (2006). Collins English Dictionary: Complete & Unabridged, Collins.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). Journal of Quantitative Linguistics, 17, 94–100.

Crystal, D. (1982). Profiling Linguistic Disability London: Arnold.

Crystal, D., Fletcher, P., & Garman, M. (1976). The Grammatical Analysis of Language Disability: A Procedure for Assessment and Remediation London: Edward Arnold.

Garside, R. (1996). The robust tagging of unrestricted text: the BNC experience. In: J. Thomas & M. Short (Eds), Using Corpora for Language Research: Studies in the Honour of Geoffrey Leech (pp. 167–180). London: Longman.

Herdan G. (1960). Type-token Mathematics. ‘s-Gravenhage: Mouton.

Herdan, G. (1966). The Advanced Theory of Language as Choice and Chance Berlin: Springer.

Hess, C. W., Ritchie, K. P., & Landry, R. G. (1984). The type-token ratio and vocabulary performance. Psychological Reports, 55, 51–57.

Hess, C. W., Haug, H. T., & Landry, R. G. (1989). The reliability of type-token ratios for the oral language of school age children. Journal of Speech, Language and Hearing Research, 32, 536.

Hughes, D. L., Fey, M. E., & Long, S. H. (1992). Developmental sentence scoring: Still useful after all these years. Topics in Language Disorders, 12, 1–12.

John, P., Brooks, B., Wand, C., & Schriever, U. (2013). Information density in bridge team communication and miscommunication – a quantitative approach to evaluate maritime communication. WMU Journal of Maritime Affairs, 12, 229–244.

Kemper, S., Kynette, D., Rash, S., O’Brien, K., & Sprott, R. (1989). Life-span changes to adults’ language: Effects of memory and genre. Applied Psycholinguistics, 10, 49–66.

Kintsch, W., & Keenan, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 5, 257–274.

Klee, T. (1985). Clinical language sampling: analysing the analyses. Child Language Teaching and Therapy, 1, 182–198.

Klee, T., & Fitzgerald, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. Journal of Child Language, 12, 251–269.

Lee, L. L., & Koenigsknecht, R. A. (1974). Developmental Sentence Analysis: A Grammatical Assessment Procedure for Speech and Language Clinicians Evanston, IL: Northwestern University Press.

Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: the tagging of the British National Corpus. Proceedings of the 15th Conference on Computational Linguistics – Volume 1 (pp. 622–628). Association for Computational Linguistics.

Levene, H. (1960). Robust Tests for Equality of Variances. In: I. Olkin (Ed), Contributions to Probability and Statistics (pp. 278–292). Palo Alto, CA: Stanford University Press.

Müller, D. (2002). Computing the type token relation from the a priori distribution of types. Journal of Quantitative Linguistics, 9, 193–214.

Richards, B. (1987). Type/token ratios: What do they really tell us? Journal of Child Language, 14, 201–209.

Rosenberg, S., & Abbeduto, L. (1987). Indicators of linguistic competence in the peer group conversational behavior of mildly retarded adults. Applied Psycholinguistics, 8, 19–32.

Scarborough, H. S. (1990). Index of productive syntax. Applied Psycholinguistics, 11, 1–22.

Turner, A., & Greene, E. (1977). The Construction and Use of a Propositional Text Base (pp. 77–63). Boulder, Colorado: Institute for the Study of Intellectual Behavior, University of Colorado.

Tweedie, F. J., & Baayen, R. H. (1998). How variable may a constant be? Measures of lexical richness in perspective. Computers and the Humanities, 32, 323–352.

Ucrel, University Centre for Computer Corpus Research on Language, Lancaster University. Retrieved 13th March 2013, from http://ucrel.lancs.ac.uk/claws5tags.html

Weitzman, M. (1971). How useful is the logarithmic type/token ratio? Journal of Linguistics, 7, 237–243.

Profiling maritime communication by non-native speakers: A quantitative comparison between the baseline and standard marine communication phraseology[1]

Peter John

Faculty of Maritime Studies, Jade University of Applied Sciences, Germany

Australian Maritime College, University of Tasmania, Australia

Benjamin Brooks

Australian Maritime College, University of Tasmania, Australia

Ulf Schriever

Australian Maritime College, University of Tasmania, Australia

Abstract

This paper compares ESP communication by non-native speakers of Maritime English with communication outside a nautical setting in order to profile its structural idiosyncrasy. Vocabulary growth, word frequencies, lexical and key word densities, and grammar diversity as dependent linguistic variables observed in transcribed full-mission simulation exercises are contrasted to the Brown Corpus, the Vienna-Oxford International Corpus of English and the Standard Marine Communication Phrases (SMCP). Using quantitative linguistics, inherent structural patterns of nautical team communication are identified and similarities and variations highlighted. Significant differences found in all linguistic features are gauged by means of the Probability of Superiority (PS) effect size. A linguistic profile is created which quantifies the observed language patterns and provides a quantitative model for the linguistic genre of this particular discourse community. The model fills the gap of quantitative research on empirical bridge team communication samples and delivers a valid tool for estimating the magnitude of observed linguistic effects.

1 The original article has been published by the ESP Journal under a Creative Commons licence: https://doi.org/10.1016/j.esp.2017.03.002

key words: Bridge team communication, Maritime English, Genre Analysis, Relative word frequencies, Maritime key words

1. Introduction

Communication on board ships has long been identified as a decisive factor for safe navigation. This importance becomes especially evident whenever a ship accident occurs, as in the disastrous evacuation procedures of passenger ships Costa Concordia and Sewol, to state two recent examples. Research has found that communication problems alone cause almost half of all marine accidents whilst miscommunication is a contributory factor in nearly all shipping accidents (for an overview, cf. John, Brooks, Wand, & Schriever, 2013; Möckel, Brenker, & Strohschneider, 2014). Most communication on board ships is verbal, but although Voyage Data Recorders (VDR) are installed on modern ships to record, amongst other information, all utterances made by navigational officers, a very limited number of authentic bridge team communication samples is available for linguistic research. This notorious scarcity of empirical information has been highlighted by Dževerdanović-Pejović (2013).

Linguistic research relies on observations on how people communicate. However, given the limited scope of authentic speech samples little quantitative research has been conducted in the domain of the bridge team communication discourse community in order to determine which speech patterns are actually used by seafarers to assess situations, carry out navigational tasks and avoid dangerous situations, and to what extent these speech acts differ from spontaneous verbal communication outside the maritime world (Cole & Trenkner, 2012; Pritchard, 2003; Trenkner, 1996; Weeks, 1997).

While it is true that authentic communication from a ship's bridge is not available in an annotated corpus for quantitative research purposes, future nautical officers are trained in full-mission ship handling simulators which replicate the navigational tasks carried out on board ship. The research presented in this paper makes use of these simulation exercises to overcome the scarcity of available on-board speech by using audio-recorded maritime simulation sessions. A verbatim transcript of the recorded communication allows corpus linguistics techniques to be used with the aim to discriminate idiosyncratic language patterns of seafarers.

1.1 Bridge team communication

Bridge team communication is a generic term for spontaneous speech acts by nautical officers who navigate the ship as a team. It shares most of the characteristics of team communication outside a nautical setting. However, given the very specific work environment in which it takes place, bridge team communication also differs regarding the team's composition, communication channels and the scope of its content. Teams on board sea-going ships are nearly always multinational and multicultural, with a substantial number of nationalities and ethnicities working very closely together and sharing one environment in which they work and live (cf. Deboo, 2004; Horck, 2005; Noble, Vangehuchten, & Van Parys, 2011). On merchant vessels, bridge teams are usually composed of the captain or shipmaster, the first, second and third navigational officer and a helmsman. Crew members work in shifts covering the ship operation twenty-four hours a day, seven days a week (cf. Jensen et al., 2006). Depending on the ship type and deployment area work shifts of four or six hours are customary. Due to this organisational scheme two to three people usually work together during their shift after which the team composition changes. One officer, who need not be the highest ranked, has the Command of Navigation (CoN) and is therefore responsible for all decisions made and actions taken during each watch.

While bridge team members engage in direct, face-to-face communication in order to assess situations and make decisions, they also communicate via UHF radio with other crew members located in different areas of the ship, e.g. in the engine room, on deck, etc., which extends the bridge team to a distributed team. Communication is also undertaken by VHF radio with the shore-based Vessel Traffic Service (VTS), tugs and other ships, and sometimes via satellite or mobile phone with the shipping company, the charterer's agents and other people ashore, so that a virtual team environment is created (for an overview on virtual team communication cf. Potter & Balthazard, 2002). If no other common language is available, the International Maritime Organisation (IMO) stipulates that crew members shall communicate in English as a lingua franca.

Given the importance of bridge team communication for the safe operation of a ship, the scarcity of publications of quantitative research in this specific discourse domain is rather surprising.

1.2 Research question and hypotheses

Observing bridge team communication in full-mission simulation opens up possibilities to gain an insight into the structure of naturally occurring language in a unique English for Specific Purposes (ESP) environment. Different speech patterns can be analysed and inferences made on their effectiveness in given situations. By contrasting maritime with non-maritime communication, similarities and differences can be singled out, and the appropriateness or idiomaticity of the language use is discernible.

The adopted research approach aims to contribute to a quantitative model of the language variety or genre of Bridge Team Communication as a sub-genre (cf. Baker & Ellece, 2011, p53) of Maritime English. By analysing a series of linguistic variables it sets out to identify and define this specific ESP variety using a descriptive approach.

These objectives lead to the following research question:

> To what extent do the speech patterns of bridge team communication by non-native speakers of English in full-mission simulation differ lexically and grammatically from other, non-nautical communication?

In order to answer this general research question the following null hypotheses are formulated:

> $H_0$1: The inter-textual vocabulary growth does not differ significantly between bridge team and other, non-nautical communication.
>
> $H_0$2: The relative word frequency distribution does not differ significantly between bridge team and other, non-nautical communication.
>
> $H_0$3: The distribution of content words does not differ significantly between bridge team and other, non-nautical communication.
>
> $H_0$4: The distribution of nautical key words does not differ significantly between bridge team and other, non-nautical communication.
>
> $H_0$5: The part-of-speech diversity distribution observed in bridge team communication does not differ significantly from other, non-nautical verbal communication.

The five hypotheses aim to profile the bridge team members' overall speech behaviour and to quantify linguistic variables which can be assumed to be idiosyncratic for the given ESP environment. $H_0$1 compares the number of different words (types) that can be expected for a given total word count (tokens). $H_0$2 studies how these word types are distributed. $H_0$3 computes differences in the corpora's lexical density. $H_0$4 identifies the distribution of specific maritime key words, and $H_0$5 highlights differences in grammar diversity.

2. Data sampling

This research compares verbatim transcripts of full-mission bridge team simulation with three different text corpora: the Brown Corpus of Standard American English (1979, 1961), the Vienna-Oxford International Corpus of English (VOICE 2013) and the SMCP text collection developed by the authors based on the Standard Marine Communication Phrases (International Maritime Organization, 2001).

2.1 Text corpora

The Brown Corpus of Standard American English (henceforth referred to as Brown Corpus) dates back to 1961 and comprises 500 texts, each of which consists of about 2,000 words. The corpus is divided into 15 different categories and contains more than one million words. In spite of it being one of the first big text corpora, it is still intensively used for linguistic research, with over five thousand publications citing it over the past ten years according to a Google Scholar search for "Brown Corpus". Given its extended use as a reference text corpus over many years, this corpus is especially apt for being used in baseline calculations. For this research, the tagged version (form C) of the revised and amplified version of 1979 has been used (Francis & Kucera, 1979).

The Vienna-Oxford International Corpus of English (henceforth referred to as Vienna Corpus) comprises in excess of one million words of naturally occurring face-to-face communication by 1,250 second language speakers with approximately 50 different first languages who use English as a lingua franca in a variety of different speech acts. For this research, the VOICE POS XML 2.0 version has been used. In order to avoid any distortions in the computed linguistic values, the tags for Breathing (BR), Laughter (LA) and Pause (PA) have been removed as these are not included in the other three corpora.

The Vienna Corpus reflects a natural usage of English as produced by non-native speakers (Seidlhofer et al., 2013).

The Brown Corpus and the Vienna Corpus are used to compute the baseline communication patterns that can be expected outside the nautical environment of bridge team communication.

The Standard Marine Communication Phrases (SMCP) were introduced by the International Maritime Organization in the year 2001 as a set of "precise, simple and unambiguous" phrases (International Maritime Organization, 2001, p3). These phrases attempt to cover all internal and external communicative situations on board sea-going ships thus reducing "problems of communication [which] may cause misunderstandings leading to dangers to the vessel, the people on board and the environment" (International Maritime Organization, 2001, p3) by using "a simplified version of maritime English in order to reduce grammatical, lexical and idiomatic varieties to a tolerable minimum, using standardized structures for the sake of its function aspects" (International Maritime Organization, 2001, p12). The "ability to use and understand the IMO SMCP is required for the certification of officers in charge of a navigational watch on ships of 500 gross tonnage or more" (International Maritime Organization, 2001, p3) and made compulsory under the International Convention on Standards of Training, Certification and Watchkeeping (STCW) for Seafarers, 1978, as revised in 1995 (International Maritime Organization, 1995). In 2010, the Manila Diplomatic Conference on the STCW Convention further strengthened the importance of effective oral communication (cf. Trenkner & Cole, 2010).

In line with its intended use, the SMCP have been published as a manual to be used in education and training. All phrases are displayed in a compacted form by using wildcard characters. In its current form the SMCP can therefore not be used for quantitative text analysis. For this reason, they have been re-worded by the authors into discrete sentences as shown in Table 1. Further examples can be found at www.smcpexamples.com (Gregorič & John 2013).

According to McEnery and Wilson (2001, p29) "any collection of more than one text can be called a corpus [...], hence a corpus may be defined as any body of text. It need imply nothing more. But the term 'corpus' when used in the context of modern linguistics

tends most frequently to have more specific connotations than this simple definition provides for". Given that the re-worded SMCP do not constitute a collection of more than one text, they are henceforth not referred to as a text corpus but as the SMCP text collection.

Original SMCP phrase	**Re-worded SMCP phrase**
A1/1.1.7.1: I am / MV … ~ not under command. ~ adrift. ~ drifting at ... knots to ... (cardinal points/half cardinal points). ~ drifting into danger.	I am not under command.
	Motor vessel Pi not under command.
	I am adrift.
	I am drifting at two knots to North North West.
	I am drifting into danger.
	Motor vessel Pi adrift.
	Motor vessel Pi drifting at two knots to North North West.
	Motor vessel Pi drifting into danger.

table 1: Original and re-worded SMCP phrases

The SMCP text collection can be considered the prescriptive language standard against which all bridge team communication can be analysed structurally and lexically. It constitutes the highest level of idiomaticity and is therefore used in this research as the reference for lexico-grammatical frequency and pattern analysis.

2.2 Bridge team transcript

The verbatim transcript of bridge team communication is based on observational data obtained from training exercises recorded in the years 2013 and 2014 at the Maritime Faculty of Jade University of Applied Sciences in Germany. The exercises included 10 bridge teams involving a total of 23 under-graduate students in their final year of Nautical Sciences who volunteered to participate in the exercises. No participants withdrew from their participation, so no attrition effects must be assumed. All students were German nationals and non-native speakers of English out of which 21 stated their mother tongue as German, one as German and Dutch and one as Tagalog. On a five-

point Likert scale (excellent, very good, good, satisfactory, poor), five students rated their English skills as very good, 14 as good and four as satisfactory (median = 3 “good”). They had worked on board sea-going ships for at least one year (median = 13 months). The sample included 22 male students and one female. Each exercise was recorded over 60 min which leads to a total recording time of 600 min. The data were collected with the informed consent of all participants involved and in compliance with the Social Sciences Human Research Ethics regulations of Jade University of Applied Sciences (Germany) and of the University of Tasmania (Australia). The transcripts were made by the first author and validated by the co-authors. Ambiguous or unintelligible words were marked with a wildcard character.

The recorded bridge team communication includes typical standard tasks carried out by navigational officers including route planning, being underway (proceeding) and assessing possible risks to navigation. Participating students communicated face-to-face with the members of their bridge team and by VHF radio with the simulated Vessel Traffic Service (VTS), the Maritime Rescue Co-ordination Centre (MRCC) and with other simulated ships. They also used UHF radio to talk to their own ship’s bosun. All radio communication partners were senior navigation officers working at the simulation facilities except for the other ships which were equipped with volunteering students. Transcript excerpt 1 includes some typical elements of direct, face-to-face communication among bridge team members.

Transcript excerpt 1

speaker	**utterance**
(1) shipmaster	What is the next course?
(2) officer	Next course in the channel is two two one.
(3) shipmaster	Two two one.
(4) officer	Uhum.
(5) officer	We are now almost abeam, erm, this one.
(6) officer	But the vessel is really hard to steer because she is rather short and the current strong and every time is, every time

	going like, like that.
(7) shipmaster	But on the other way it is, er, she's very high manoeuvrable.
(8) officer	Yes, we have six zero rudder angle, yeah, what do you expect?
(9) shipmaster	But if you, but if you use six zero it will be really difficult to...
(10) officer	Oh, we really don't have to use six for the moment.
(11) shipmaster	She's coming very, very quick.
(12) officer	Yeah, the course is two two one.
(13) shipmaster	Hm.
(14) shipmaster	But now I have really to take a close look because there is one vessel coming here.
(15) officer	But if you look, have a look here.
(16) shipmaster	Yeah, oh, this one.
(17) shipmaster	Yeah, she's turning very good, so I would turn then really hard.
(18) officer	Yeah, from this buoy to this buoy the channel is a straight line, so.
(19) shipmaster	Okay, yeah, no problem.
(20) officer	Yeah.

Transcript excerpt 1 clearly illustrates how the shipmaster and nautical officer discuss a possible risk, develop a shared mental model and agree on measures to be taken. The excerpt includes typical elements of verbal communication and some elements of its specific ESP context.

3 Data analysis

In the analysis of the collected data the following steps are undertaken for accepting or rejecting each null hypothesis: Firstly, the reasoning for the hypothesis and descriptive statistical information are presented for the linguistic feature to be analysed. Secondly, the analytical method is presented and pre-requisites are defined and tested. Thirdly, the analysis is carried out, and fourthly, findings are summarised.

	Baseline		**Empirical data**	**Upper limit**
	Vienna Corpus	Brown Corpus	Bridge team transcript	SMCP Corpus
word tokens	1,016,399	1,036,125	43,019	46,529
word types	17,449 ratio 0.017	40,187 ratio 0.039	1,843 ratio 0.043	1,883 ratio 0.040
content words	394,230 ratio 0.387	558,066 ratio 0.538	20,767 ratio 0.482	31,879 ratio 0.685
adjectives	49,535 ratio 0.048	71,994 ratio 0.069	1,559 ratio 0.036	2,426 ratio 0.052
adverbs	79,155 ratio 0.077	37,898 ratio 0.036	4,911 ratio 0.114	1,183 ratio 0.025
nouns	155,897 ratio 0.153	270,978 ratio 0.261	7,592 ratio 0.176	17,846 ratio 0.383
verbs	92,619 ratio 0.091	115,842 ratio 0.111	3,819 ratio 0.088	4,878 ratio 0.104
numerals	61,354 ratio 0.060	17,024 ratio 0.016	2,886 ratio 0.067	5,546 ratio 0.119
function words	622,169 ratio 0.613	478,059 ratio 0.462	22,252 ratio 0.518	14,650 ratio 0.315
key words	189,397 ratio 0.186	208,633 ratio 0.020	13,212 ratio 0.307	25,601 ratio 0.550

table 2: Overview of word count and ratios in analysed text corpora

Table 2 displays word counts for the Brown and Vienna Corpus, the Bridge Team transcript and the SMCP text collection. Given the substantial differences in the number of words each text corpus contains, it also states ratios for the given subcategories, e.g. the ratio of different word classes to the total word count.

Hypothesis $H_0$1: The inter-textual vocabulary growth does not differ significantly between bridge team and other, non-nautical communication.

Vocabulary growth describes the changing relation of word types (vocabulary size) to word tokens (total word count) over an increasing text length (inter-textual growth) or a defined time frame. It has been extensively used for estimating lexical diversity in first and second language learners (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Laufer, 1998; Nagy & Scott, 2000; Verhoeven, van Leeuwe, & Vermeer, 2011) where it is assumed that an increase in vocabulary growth correlates with an advancing language learning process.

In maritime communication, a higher lexical diversity does not necessarily lead to an improved communication by its discourse community. To the contrary, the standard phraseology was introduced as a coded language which reduces lexical richness on purpose to remove any ambiguities and provide a simple and clear language. Table 2 hints at differences between the type-token ratios (TTR) with figures ranging from 0.017 for the Vienna Corpus, 0.039 for the Brown Corpus, 0.043 for the Bridge Team transcript and 0.040 for the SMCP text collection. However, these figures are biased as the TTR was found to differ in relation to the length of the chosen text samples (cf. Covington & McFall, 2010). For this reason, the corpora were split into samples of 43,019 words which is the exact length of the Bridge Team transcript. Following this method, the Brown and Vienna Corpus were both divided into 23 samples totalling 989,437 words in each corpus. In both corpora, the remaining words were disregarded. To take full advantage of the much smaller SMCP text collection, 10 random samples of 43,019 words each were drawn up without removing the extracted samples.

Figure 1 displays the number of text types for the text token chunks in the three corpora.

The SMCP text collection presents a type-token ratio which is much closer to that of the Bridge Team transcript, with TTR values of 0.040 and 0.043, respectively. The expected number of types for any given number of tokens (up to a value of 43,019) can be

calculated by the power functions $V(n)_{SMCP}=0.99n^{1.44}$ (SMCP, $R^2=0.989$) and $V(n)_{BT}=0.02n^{1.91}$ (Bridge Team transcript, $R^2=0.995$).

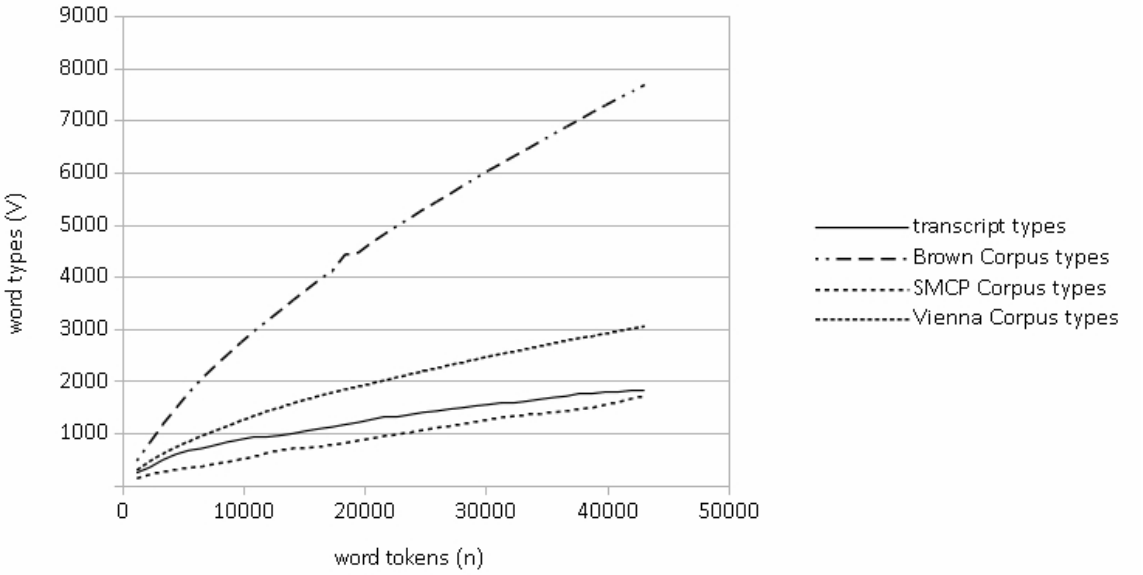

figure 1: Inter-textual vocabulary growth in text corpora for chunks of 43,019 words

Although it seems obvious that the distribution of the type-token ratio differs clearly, at least between the Bridge Team transcript and the Brown and Vienna Corpus, the four corpora were also tested statistically in order to quantify the significance of the findings, i.e. the probability for the effects to occur by chance.

To be able to compare the samples' variances, these have to be tested first for a normal distribution as this determines which statistical methods can be employed. Anderson–Darling's test was performed for testing the corpora's distribution against a normal distribution, leading to $p=0.25$ for the Brown Corpus, $p=0.19$ for the Vienna Corpus, $p=0.37$ for the SMCP text collection and $p=0.12$ for the Bridge Team transcript. All samples are therefore assumed to be normally distributed.

The text corpora were also tested for homoscedasticity employing Levene's test for homogeneity of variances. For the four corpora together it resulted in $p<0.000$ for the Brown and the Vienna Corpus and $p=0.76$ for the SMCP Corpus so that a homoscedastic distribution can only be assumed between the Bridge Team transcript and the SMCP text collection.

Given the text corpora's heteroscedasticity, non-parametric statistical testing methods were chosen. As figure 1 clearly illustrates, the different word type samples (i.e. the different text corpora) do not originate from the same distribution. The non-parametric Kruskal-Wallis one-way analysis of variance carried out on the four text corpora corroborates these findings with $p<0.000$ (two-tailed, $\alpha=0.01$, confidence level 99%). Mann–Whitney–Wilcoxon tests were performed post-hoc between the Bridge Team transcript and the other three text corpora. For the SMCP it resulted in $p=0.005$ and for the Brown and the Vienna Corpus it resulted in $p<0.000$ (two-tailed, $\alpha=0.01$, confidence level 99%). The null hypothesis $H_0 1$ must be rejected as the differences in the inter-textual vocabulary growth are highly significant between the four text corpora observed[1].

The vocabulary growth observed in the Bridge Team communication sample differs very significantly from the Brown and Vienna Corpus. Bridge team communication by non-native speakers uses a far more restricted vocabulary size than that observed in written English and, more importantly, than in verbal communication by non-native speakers outside a maritime setting. In comparison with the Brown Corpus, non-native bridge team members can be expected to use between 40 and 50% of word types for any text length of up to 5,000 words, 30-39% for a text length of less than 12,000 words and 29-24% for a text length of up to 43,000 words. For the Vienna Corpus the percentages are 70 to 80% in texts of less than 5,000 words and 60 to 79% in texts of less than 43,000 words.

Compared with the SMCP text collection, the Bridge Team communication sample contains only 1-1.9% more types for any given text size up to 43,000 words.

Hypothesis $H_0 2$: The relative word frequency distribution does not differ significantly between bridge team and other, non-nautical communication.

Word frequency distributions provide information on the lexical proximity of texts. Word types that appear at a similar frequency indicate a relatedness of two texts. On the other hand, types that appear more frequently in one text than in another can be considered key words of a specific field covered by the given text.

The frequency profiling technique detects differences of relative word frequencies between any two texts. It is independent from a text's length, so that the Brown and

Vienna Corpus could be used without splitting them into smaller fragments. A frequency list of the 406 word types that occurred in all of the four text corpora was produced and the log likelihood statistic (LL) was calculated using a contingency table as suggested by Rayson and Garside (2000). Subsequently, the expected frequency values and the log likelihood statistic itself were computed for each word type. Table 3 presents a selection of the calculated values. The plus sign describes a higher frequency and the minus sign denotes a lower frequency of the word type in the respective text corpus compared to the Bridge Team transcript.

word type	**LL Bridge Team transcript – Vienna Corpus**	**LL Bridge Team transcript – Brown Corpus**	**LL Bridge Team transcript – SMCP text collection**
ahead	+1.7	-86.1	-86.1
alarm	+24.6	-47.8	-46.7
anchor	+7.5	-19.6	+11.5
bridge	+0.7	-30.9	-34.3
depth	0.0	-29.2	-27.0
lead	+11.4	-6.7	-7.4
message	+2.4	-4.2	-2.1
navigation	+3.8	-17.4	-14.1
same	+8.7	+37.0	+73.8
speak	+0.8	+0.3	+65.0
within	+116.8	+10.9	+13.3

table 3: Log likelihood statistics for the analysed text corpora

An LL value of zero means that a given word type appears in both text corpora at an identical frequency, and the higher the LL value, the more significant are the differences in the relative frequencies of the two text corpora. Adopting a significance level of a = 0.01, the null hypothesis $H_0 2$ must be rejected for LL values equal or higher than the critical value of 6.63. Following this procedure for all 406 word types it was found that

in comparison with the SMCP text collection, 224 word types presented a significantly different frequency. For the Brown Corpus this value amounted to 305 word types and for the Vienna Corpus to 292. Given the fact that out of a total of 1,843 word types observed in the Bridge Team communication only 406 types were tested for their relative frequencies in the other three text corpora while the remaining 1,383 word types do not appear at all in at least one of the other corpora, $H_0$2 is rejected.

In his paper "Language is never, ever, ever, random", Kilgarriff (2005, p263) points out that "[l]anguage users never choose words randomly, and language is essentially non-random". This non-randomness of language is clearly reflected in relative word frequencies where a log likelihood statistic of zero for all word types is only possible in two identical texts. However, when studying the LL values beyond the mere hypothesis testing, they are indicative of the divergence between the different text corpora. The median LL value for the 406 words shared by the Bridge Team communication and the Brown Corpus is 18.92 while for the Vienna Corpus it is 15.09. As in the case of inter-textual vocabulary growth, the relative word frequencies are more closely related to the utterances by non-native speakers included in the Vienna Corpus than to the more formal English written by native speakers as reflected in the Brown Corpus.

The LL statistic for the comparison with the SMCP text collection is 8.73, a value which is expectedly much closer to the Bridge Team communication than the other, more general text corpora. Here, 44% of the compared word types occurred at a similar relative frequency as opposed to only 25% in the case of the Brown Corpus and 28% in the Vienna Corpus.

Hypothesis $H_0$3: The distribution of content words does not differ significantly between bridge team and other, non-nautical communication.

Words can be differentiated according to their grammatical and lexical function, leading to the lexical density concept which considers the ratio of content words to the total word count (cf. Halliday, 1987). Content words include nouns, verbs, adjectives, adverbs and numerals whereas function words include all other word classes. By computing the lexical density of a given text, inferences can be made about its information content. This is important as communication by work teams mostly aims at interchanging information. The more information is exchanged effectively (i.e. correctly transmitted,

received and understood), the higher the linguistic effectiveness of speech acts. In a safety-critical environment like a ship's bridge, this can be a decisive factor. As an example, in transcript excerpt 2 all content words are underlined and the lexical density (LD) is given for each utterance.

Transcript excerpt 2

speaker	utterance	LD
(21) shipmaster	I will keep it now with fifty percent and, er.	0.40
(22) shipmaster	When we cross all these vessels, then, I'm happy.	0.40
(23) officer	But no more traffic is out there.	0.57
(24) shipmaster	We really pass, er?	0.50
(25) officer	Like this, what's, what's that on the left side?	0.18
(26) shipmaster	There's one more vessel, I see already.	0.50
(27) shipmaster	Okay, anyway, we'll check this with the radar.	0.44
(28) shipmaster	Ha, ha, ha, uh, okay, now I see the buoys.	0.40
(29) shipmaster	This is very nice, I can mark them already.	0.44
(30) shipmaster	There is one more vessel but it shouldn't be a problem.	0.25
(31) shipmaster	Okay, now I will go starboard again, yeah?	0.63
(32) officer	Yeah.	0.00

This example illustrates the substantial differences in lexical density. The highest value observed in (31) clearly contains more information than (25) or (32). The word count and ratios for content and function words in the text corpora are presented in Table 2. Figure 2 compares the different content word classes observed.

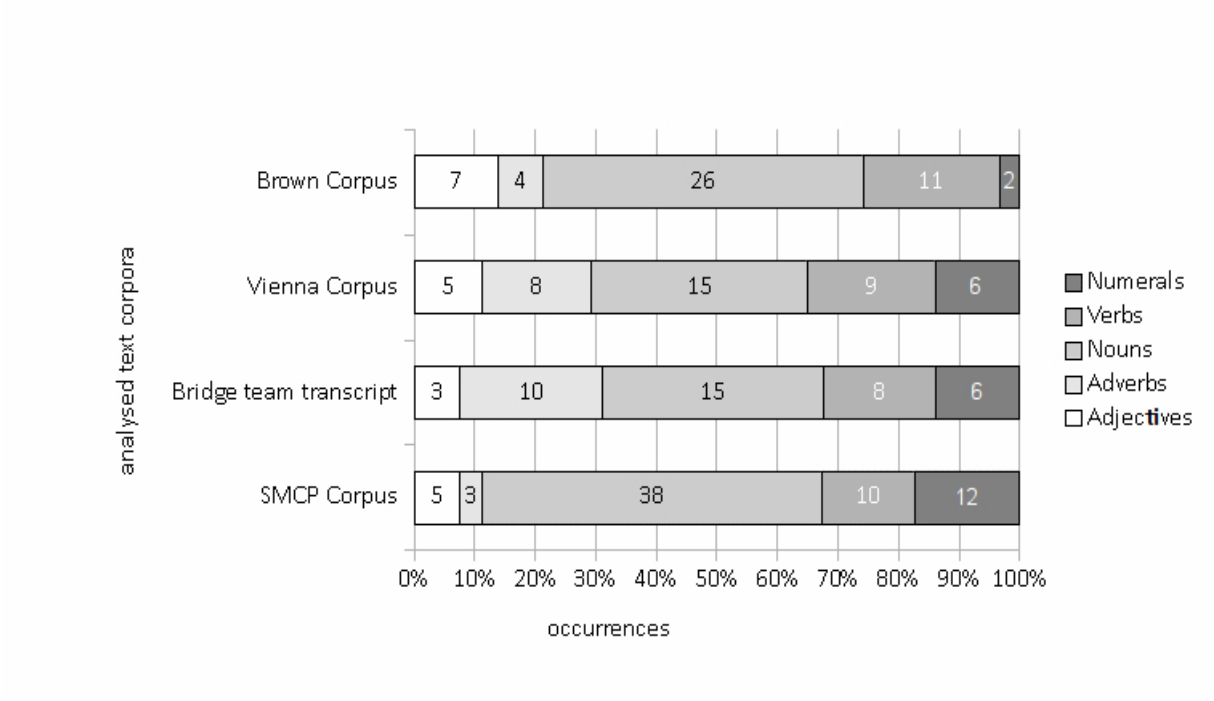

figure 2: Occurrences of content word classes in text corpora

In the following, the probability for the differences displayed in Figure 2 to occur randomly is computed for all content words together and for each individual content word class. If the calculation was simply performed on the counted words of a particular class in each utterance, this would obviously not take into account the differences in the utterance length (word count per each utterance). For this reason the calculated ratio of a particular word class is considered for any given utterance. So, if an utterance with a total length of ten words contains four content words, its content word ratio of 0.40 is used. If the same utterance contains two adjectives, its adjective ratio of 0.20 is used for the computation.

As figure 3 well illustrates, the different content word samples (i.e. the different text corpora) do not originate from the same distribution. The non-parametric Kruskal-Wallis one-way analysis of variance carried out on the four text corpora corroborates these findings with $p<0.000$ (two-tailed, $\alpha=0.01$). Mann–Whitney–Wilcoxon tests were performed post-hoc between the Bridge Team transcript and the other three text corpora. For the SMCP and the Vienna Corpus it resulted in $p<0.000$, and the Brown Corpus in $p=0.001$ (two-tailed, $\alpha=0.01$).

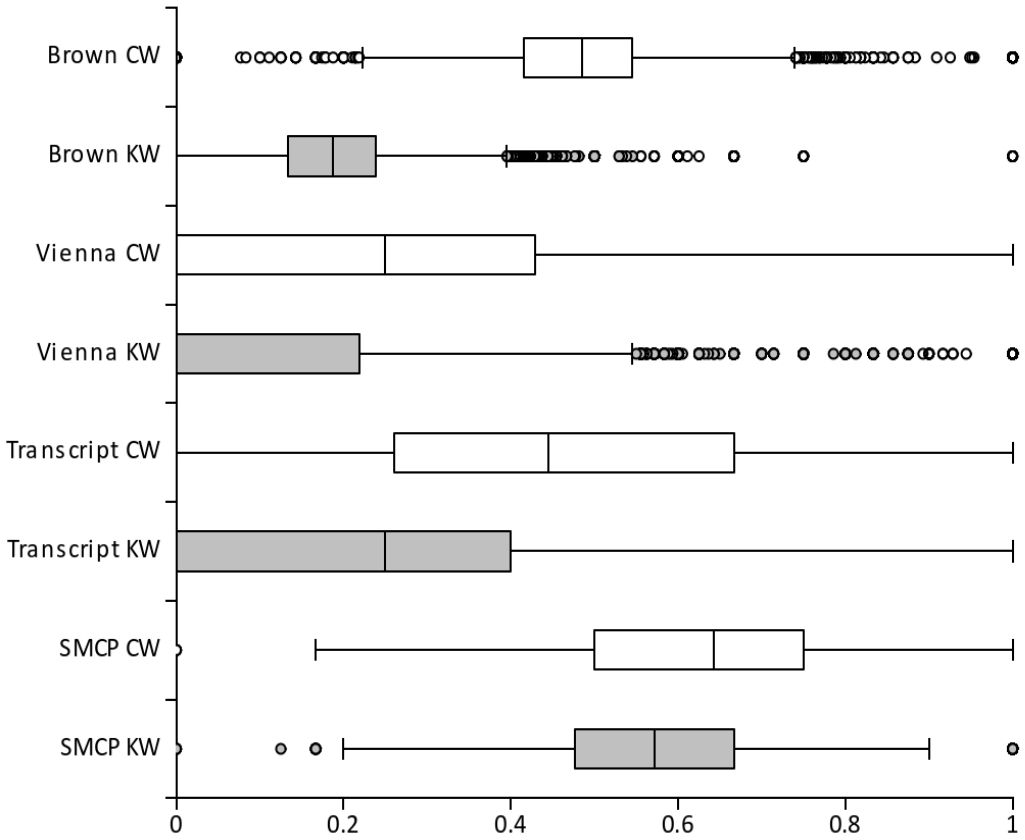

figure 3: Content word and key word distribution in analysed text corpora

The Kruskal–Wallis test carried out for occurrences of verbs, nouns, adjectives, adverbs and numerals in the different text corpora leads to $p<0.000$, and post-hoc Mann–Whitney U tests on the five word classes between the individual corpora also resulted in $p<0.000$ for all combinations. In all tests, the significance level was set to $\alpha=0.01$.

In order to estimate the strength of the differences between the individual content word ratios, the effect size is computed as their Probability of Superiority (PS) which does neither assume a normal nor a homoscedastic distribution of the observed variables. Grissom and Kim (2005, p98) have described this effect size as "the probability that a randomly sampled member of population a will have a score (Ya) that is higher than the score (Yb) attained by a randomly sampled member of population b."

The effect sizes between the Bridge Team transcript and the other corpora are displayed in Table 4.

PS	content words	adjectives	adverbs	nouns	verbs	numerals
Brown Corpus	0.49	0.28	0.54	0.32	0.33	0.44
Vienna Corpus	0.70	0.49	0.61	0.59	0.54	0.58
SMCP text collection	0.30	0.41	0.65	0.18	0.34	0.46

fable 4: Effect sizes of content word classes

$H_0$3 must be rejected because the distribution of content and function words in bridge team communication differs significantly from other, non-nautical communication. Bridge team communication presents a higher lexical density than the non-nautical, verbal communication collected in the Vienna Corpus (VC). The distribution between both corpora differs strongly with PS = 0.70. The lexical density of bridge teams is nearly identical with the written texts in the Brown Corpus (BC) with PS = 0.49. The medians in both text corpora are relatively close (transcript = 0.44, BC = 0.50 as opposed to VC = 0.25). However, the transcript's lexical density varies more markedly than that of the Brown Corpus, which is reflected in its wider inter-quartile range (transcript 0.26–0.76; BC 0.42–0.55). The observed bridge teams' lexical density differs starkly from that of the SMCP text collection, with PS = 0.30. Had the bridge team only used the SMCP, the effect size would have been 0.50.

Looking at the different content word classes, strong differences between the bridge teams and non-native speakers in a non-nautical setting were found in the distribution of nouns, clear differences in the use of verbs, adjectives and adverbs, and slight differences in numerals. Compared to written texts, bridge teams can be expected to produce far more adverbs and numerals, but fewer nouns, verbs and adjectives.

The SMCP text collection has a strikingly higher proportion of nouns with an effect size of PS = 0.18 and a distinctly higher proportion of verbs. It clearly contains more adverbs and fewer adjectives. The distribution of numerals is nearly identical. In the SMCP, the ratio of nouns is more than twice as high as the corresponding value in the Bridge Team

transcript. The unambiguous nature of the SMCP wording is, among others, also reflected in the preference of nouns over determiners such as this, that or another. In the SMCP text collection, only 160 determiners are used which corresponds to a ratio of 0.003, whereas in the Bridge Team transcript, 1,695 determiners were produced, equalling a ratio of 0.039, a value more than ten times higher. Calculating the ratio for nouns and determiners together in both text corpora, the difference is still high (transcript = 0.215, SC = 0.386), but a tendency can be identified to replace nouns with determiners in the transcribed verbal communication. Transcript excerpt 3 offers an example (appropriate SMCP wording is given in brackets).

Transcript excerpt 3

speaker	utterance
(1) officer	No, no, these, these, these two are cardinal buoys. *(instead of: We are passing cardinal buoys H and F on port side.)*
(2) shipmaster	Yes.
(3) officer	Then this one is a...
(4) shipmaster	Er, tower, whatever.
(5) officer	Yeah, but this one is a, is a target. *(instead of: Sailing boat Alpha is a target.)*
(6) shipmaster	Yeah, but you showed here, so I was confused.
(7) officer	I have such big fingers, so it's very confusing.
(8) shipmaster	I think this, this must be the sailing vessel. (instead of: The target must be sailing boat Alpha.)
(9) officer	Yeah, this one's sailing boat. (instead of: Yes, the target is sailing boat Alpha.)

Hypothesis $H_0$4: The distribution of nautical key words does not differ significantly between bridge team and other, non-nautical communication.

Based on the differences observed in the distribution of content words it was found that utterances by bridge team members present a higher information density than that of speakers transcribed in the Vienna Corpus. However, these findings alone are not suitable to test assumptions related to their appropriateness or maritime idiomaticity. For this reason, a set of key words has been established by isolating all content words included in the SMCP (cf. John et al., 2013). As this method does not include any qualitative differentiation, it does not discriminate between words which are exclusively used in the maritime domain (e.g. anchor, starboard, leeward), those also used outside a maritime setting (e.g. bow, proceed, target) and those without any specific maritime meaning (e.g. dangerous, under, way). Interestingly, this quantitative approach for extracting key words has resulted in a very similar key word ratio (~31%) to the "coverage" reported by Chung and Nation (2003) for anatomy texts. In the following, the empirical data are compared with the two non-nautical text corpora and the SMCP text collection. It is assumed that a higher key word density reflects a higher degree of maritime idiomaticity.

Table 2 shows that the Brown and Vienna Corpus feature a key word ratio of 0.020 and 0.186, respectively. Neither of the two corpora contains specifically maritime texts. The calculated values must therefore be assumed to be the noise generated by the adopted methodology. It is interesting to see that in verbal communication by non-native speakers the proportion of nautical key words is much lower than in texts written by native speakers. This may be explained by the fact that only a small number of the key words are used in normal conversation. To proceed, for example, is more likely to appear in written texts than in verbal communication.

The transcribed bridge team communication leads to a key word ratio of 0.307, i.e. roughly one out of three words uttered is a key word. Had the bridge team only used the SMCP, they would have produced a ratio of 0.550 as reflected in the SMCP text collection. Figure 3 displays the key word ratio for the different text corpora as box plots with the inter-quartile ranges in grey (labelled "KW"). The differences in the distribution of the SMCP content words and SMCP key words are caused by the high

number of proper names (e.g. for ships, buoys, etc.) which are counted as content words but not as key words.

In order to perform an analysis of the samples' variances, they are tested for a normal and for a homoscedastic distribution. Anderson–Darling's test for normal distribution on each of the four text corpora and Levene's test for homogeneity both resulted in $p < 0.000$, so that neither a normal nor a homoscedastic distribution can be assumed. Again, the non-parametric Kruskal–Wallis test was employed on the four text corpora resulting in $p < 0.000$. Post-hoc Mann–Whitney U tests between the Bridge Team transcript and the other three text corpora also lead to $p < 0.000$ for each of them.

For the key word distributions, the following effect sizes were computed:

$PS_{KW:Transcript/BC}=0.59$

$PS_{KW:Transcript/VC}=0.68$

$PS_{KW:Transcript/SCMP}=0.13$

$H_0 4$ must be rejected and the alternative hypothesis accepted as the key word distribution between the text corpora differs significantly between the four text corpora. The most striking difference here is not the strong effect sizes between the Bridge Team transcript and the two non-nautical corpora because a much higher proportion of nautical key words is expected in the Bridge Team transcript. It is rather the very strong effect size between the transcript and the SMCP text collection of 0.13. Expressed in real figures this means that in the recorded bridge team communication a high proportion of 2,064 utterances do not contain any key word at all, which equals a ratio of 0.047. The key word median, which covers 50% of all utterances by the bridge team members, equals 0.25, a value which only about 3% of the phrases in the SMCP text collection present. The first three quartiles or 75% of all utterances result in a value of 0.40 which does not even reach the SMCP text collection's first quartile of 0.48. Transcript excerpt 4 and 5 present two speech acts with different key word ratios. Transcript excerpt 4 contains a total of 44 words out of which nine are key words, the key word ratio equalling 0.20. SMCP key words are underlined.

Transcript excerpt 4

	speaker	utterance
(10)	shipmaster	What is this? But this is no, no anchorage.
(11)	officer	Maybe the wrong chart?
(12)	shipmaster	Yeah, but also.
(13)	officer	Where is it here? Where is this, where is this anchorage?
(14)	shipmaster	Yeah, this, this one, or? Neue Weser Reede, maybe? This would be South of Helgoland, but I...

Excerpt 5 consists of 30 words out of which ten are key words leading to a key word ratio of 0.33 which is slightly higher than the observed ratio in the whole transcript.

Transcript excerpt 5

	speaker	utterance
(15)	shipmaster	What did you say what was the distance to this one? To make the...
(16)	officer	Five cables. And then we have to go to the new course.
(17)	shipmaster	Hm, okay. Steady.
(18)	officer	Steady.

Hypothesis $H_0$5: The grammar diversity distribution observed in bridge team communication does not differ significantly from other, non-nautical verbal communication.

So far, analyses have been performed on lexical structures including vocabulary growth and word frequencies as well as lexical and key word densities. This last hypothesis deals with possible differences in grammatical diversity between the utterances by bridge team members, the Vienna Corpus and the SMCP text collection. The Bridge

Team transcript is also compared to verbatim transcripts of the radio programme Lingua Franca in which a presenter interviews guests, a programme which has been analysed by John and Brooks (2014). For this analysis, the Brown Corpus is disregarded, because here the focus lies exclusively on spoken discourse.

Grammar diversity has been found to vary in dependence on the utterances' length. For this reason, the specific grammar diversity index (spdi) has been developed to determine "the relative deviation of the grammar diversity value in an observed utterance from the expected number given by the PDI [POS Diversity Index] for a specific utterance length" (John and Brooks, 2014, p33).

The grammar diversity displayed as spdi box plots in Figure 4 is more homogeneously distributed than the four lexical distributions analysed above. The Bridge Team transcript's inter-quartile range stretches from 0.83 to 1.05 which is slightly lower than that of the Lingua Franca transcript (0.94–1.12), relatively similar to the Vienna Corpus (0.88–1.02) and slightly higher than that of the SMCP text collection (0.75–1.05). In the Bridge Team transcript and the Vienna Corpus, the median equals 1.00 whereas it is 0.92 in the SMCP text collection and 1.04 in the Lingua Franca radio programme.

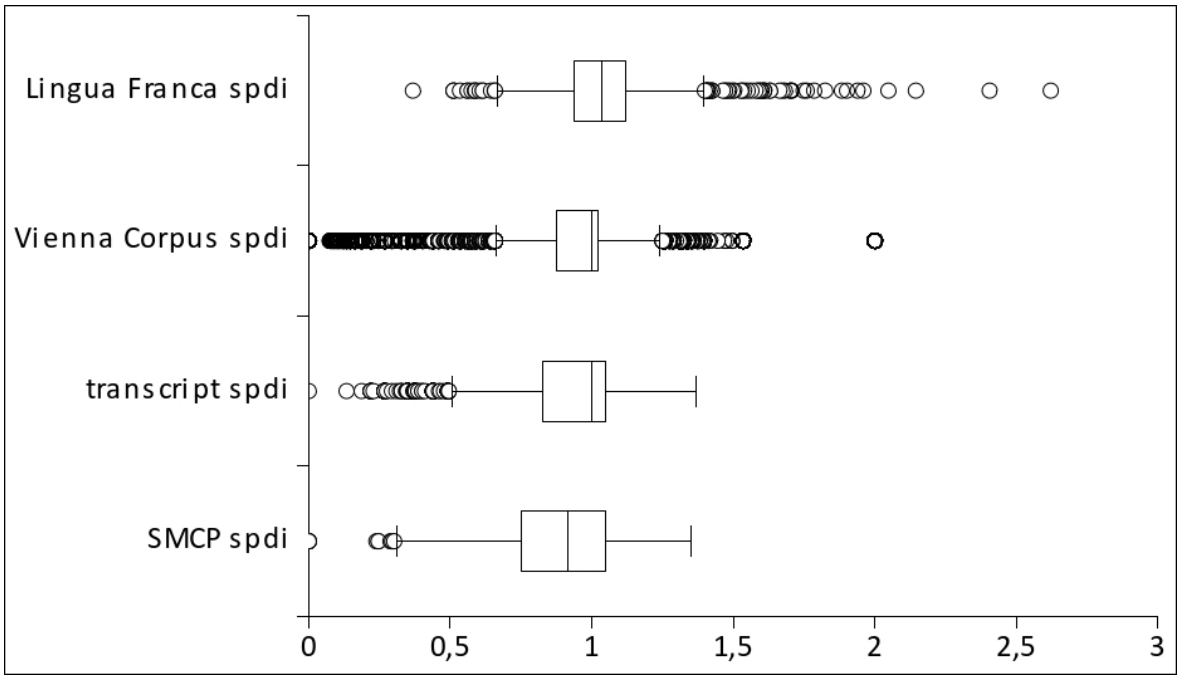

figure 4: Special POS diversity index for analysed text corpora

A normal distribution was only found for the Lingua Franca transcript. The other two text corpora and the SMCP text collection do not present a normal distribution. Levene's

test for homogeneity led to $p < 0.000$, so that no homoscedasticity can be assumed. The Kruskal–Wallis test again resulted in $p < 0.000$. The three Mann–Whitney U tests performed post-hoc between the Bridge Team transcript and the other text corpora also resulted in $p < 0.000$; the null hypothesis H 0 5 is therefore rejected.

For the spdi distributions, the following effect sizes were computed:

$$Ps_{spdi:Transcript/LF}=0.35$$

$$Ps_{spdi:Transcript/VC}=0.65$$

$$Ps_{spdi:Transcript/SCMP}=0.56$$

The grammar diversity in bridge team communication by non-native speakers resembles the speech acts by non-native speakers outside a maritime setting as transcribed in the Vienna Corpus. Although the statistical testing has resulted in a rejection of the null hypothesis, the small effect size between the two non-native speaker's corpora of PS = 0.65 nevertheless indicates a close proximity of the observed grammar structures. The grammar structures in the SMCP text collection presents an even higher degree of similarity with PS = 0.56. To the contrary, the transcripts of the Lingua Franca radio programme differ more strongly with PS = 0.35 pointing at more elaborate speech acts.

4. Limitations of the adopted methodology

No research design is totally free of a certain degree of limitations and biases. This research studies transcribed verbal communication while disregarding any para-verbal (speed, intonation, etc.) and non-verbal communication (eye movement, facial expression, gestures, etc.). It is based on simulation sessions carried out with volunteering university students of Nautical Sciences. Here, a general symmetry in speech facilitates an ideal speech situation in which bridge team members participate under an assumption of equality (Habermas, 1979) which will not always be the case on board a sea-going ship. Another difference occurs due to the fact that only one team was inter-cultural while the other bridge teams were all mono-cultural. The introduced sampling bias had to be conceded by the authors because a data sampling in real-life conditions is virtually impossible.

With regards to the samples' representativeness, the bridge team communication sampled from the German teams can only be considered representative of a German

group of prospective nautical officers with a prolonged English language learning experience. However, the authors would like to argue that this group can also be considered representative for all northern European non-native speakers of English with a similar exposure in school and university education.

5. Findings and discussion

This research identifies lexical differences and similarities of quantitative linguistic features between the specific ESP setting of bridge team communication and other text corpora in order to profile a non-native bridge team's idiosyncratic speech patterns. For this purpose, five null hypotheses were tested for significant differences, and inferences were made to quantify the observed statistics.

It was found that the inter-textual vocabulary growth ($H_0$1) of bridge team communication is not only significantly gentler than in written English but also more gradual than in spoken English outside a maritime setting. In a transcript consisting of 43,019 tokens, only 1,843 types can be expected, i.e. a type-token ratio (TTR) of 0.042 (Brown Corpus=0.179, Vienna Corpus=0.071). For any given number of tokens (n) up to 43,019, the corresponding number of types (V) can be computed by using the power function $V(n)BT=0.02^{1.91}$ (R^2=0.995). A comparison with the SMCP text collection validated the similarities of the natural speech patterns with the mandatory, coded SMCP language with its 1,883 word types and a type-token ratio of 0.040.

Relative word frequencies ($H_0$2) were also found to differ significantly between bridge team communication and the other analysed text corpora. The four corpora only shared 406 word types of which the Brown Corpus included 305 with a significantly different distribution. Again, this Corpus presented the highest difference with a median log likelihood statistic of 18.92 and only 21% of all shared word types presenting a similar frequency. For the Vienna Corpus, the corresponding values were 15.09 and 28%. The bridge team transcript is much closer again to the SMCP text collection with a median log likelihood statistic of 8.73 and 44% of word types occurring at a similar relative frequency.

A significantly differing distribution of content words ($H_0$3) was ascertained between the Bridge Team transcript and the other corpora. In total, the content word ratio of the speech acts by the bridge team members was more similar to the written English

reflected in the Brown Corpus, with an effect size of PS=0.49 as opposed to PS=0.70 for the Vienna Corpus and PS=0.30 for the SMCP text collection.

Nautical key words ($H_0$4) are assumed to indicate the appropriateness or idiomaticity of bridge team communication. The transcript's key word ratio of 0.307 is smaller than that of the SMCP text collection (0.550) but higher than in the Brown and Vienna Corpus (0.020 and 0.186, respectively). Nevertheless, the bridge teams' key word distribution differed significantly from the other corpora and surprisingly the strongest effect size was observed in the comparison with the SMCP text collection, with PS=0.13.

To compare the Bridge Team transcript's special POS diversity index ($H_0$5), the Brown Corpus was replaced by the transcript of the Lingua Franca radio programme analysed by John & Brooks (2013). A significantly different distribution was found between the corpora although the effect sizes between the Bridge Team Transcript and the SMCP text collection were relatively small in comparison to the Vienna Corpus and Lingua Franca programme.

So far, the dependent linguistic variables have been analysed individually. However, by combining them, a linguistic profile is created which effectively quantifies the observed language patterns as a coherent whole. This profile is presented graphically in figure 5 as a five-pointed polygon in which the variables have been normalised for a clear visualisation. The overlapping areas underline the much closer proximity between the Bridge Team transcript and the SMCP text collection than to the other two text corpora. The figure also shows that the observed bridge team communication still differs clearly from the mandatory coded language of the Standard Marine Communication Phrases.

Profiling the inherent communication patterns of the collected speech samples leads to a quantitative model for the specific discourse community of bridge team members in full-mission simulation exercises speaking English as their second language. This research has shown that the created model is a valid tool for quantifying the differences and similarities between the genre of bridge team communication and other communicative settings.

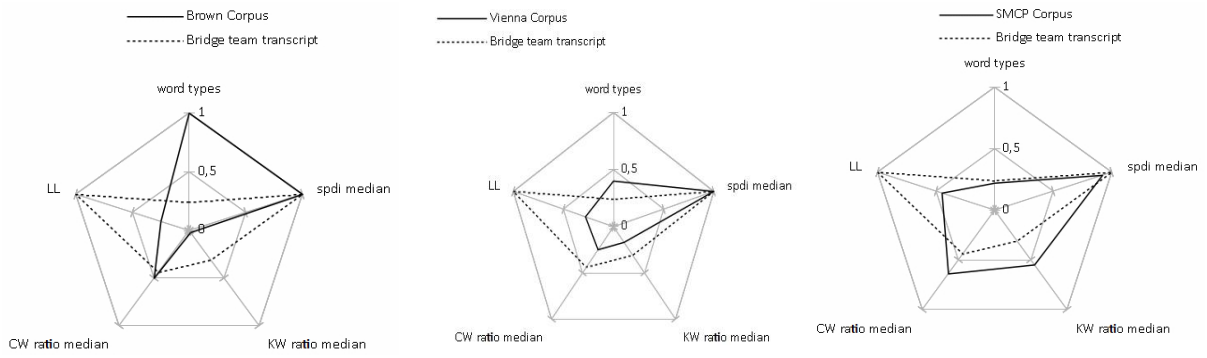

figure 5: Linguistic profile comparison of the analysed text corpora

Future research should profile the idiosyncratic language patterns of differently composed bridge teams (e.g., native English speakers) for comparison with the profile represented in this paper. Such a comparison will identify differences in the discourse between non-native and native English speakers in this particular environment, and those results may be generalisable to work teams in other domains. Correlating the profile with other behavioural data that identifies the quality of the communication may also be useful. Finally, this may lead to a model of standard communication behaviour which can be used as a benchmark in the training of future nautical officers. Communication is a decisive component of social interaction and a contributory factor to improve safety at work. In shipping, improved education and training in bridge team communication will take us a step closer towards avoiding fatal accidents as in the case of passenger vessels Costa Concordia, Sewol and many others.

References

Baker, p. & Ellece, S. (2011). *Key Terms in Discourse Analysis*. London: Continuum.

Chung, T. M. & Nation, p. (2003). Technical vocabulary in specialised texts. *Reading in a foreign language*, 15(2), 103-116.

Cole, C., & Trenkner, p. (2012). Whither Maritime English? - 2012. *Proceedings of the International Maritime English Conference* (IMEC). Yangong: Maritime Lecturers' Association, 3-18.

Covington, M., & McFall, J. D. (2010). Cutting the gordian knot: The Moving-Average Type-Token ratio MATTR. *Journal of Quantitative Linguistics*, 172, 94-100.

Deboo, K.M. (2004). Maritime Resource Management online. Alert. Internet resource available at http://www.he-alert.org, viewed on 18 April 2014.

Dževerdanović-Pejović, M. (2013). Discourse of VHF Communication at Sea and the Intercultural Aspect. *International Journal for Traffic and Transport Engineering*, 34, 377-396.

Francis, W. N. & Kucera H. (1979). Brown Corpus: A Standard Corpus of Present-Day Edited American English, for use with Digital Computers. *Department of Linguistics, Brown University. Revised 1971, Revised and Amplified 1979*. Providence: Rhode Island.

Gregorič, T. & John, p. (2013). SMCPexamples – Learn SMCP Phrases by Examples. Internet resource available at http://www.smcpexamples.com, viewed on 30 June 2014.

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research. Univariate and Multivariate Applications*. Hove: Taylor & Francis.

Habermas, J. (1979). *Communication and the Evolution of Society*. Vol. 29. Boston: Beacon Press.

Halliday, M.A.K. (1987). Spoken and written modes of meaning. In R. Horowitz & S.J. Samuels (Eds). *Comprehending Oral and Written Language*. San Diego: Academic, 55-82.

Horck, J. (2005). Getting the best from multi-cultural manning. *BIMCO Bulletin*, 100(4), London: Baltic and International Maritime Council, 28-36.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental psychology*, 27(2), 236-248.

International Maritime Organization (1995). *Standards of training, watchkeeping for seafarers, as amended*. London: International Maritime Organization.

International Maritime Organization (2001). *Resolution A.91822: IMO Standard Marine Communication Phrases*, London: International Maritime Organization.

Jensen, O. C., Sørensen, J. F., Thomas, M., Canals, M. L., Nikolic, N., & Hu, Y. (2006). Working conditions in international seafaring. *Occupational medicine*, 566, 393-397.

John, P., & Brooks, B. (2014). Lingua Franca and its Grammar Footprint: Introducing an Index for Quantifying Grammatical Diversity in Written and Spoken Language. *Journal of Quantitative Linguistics*, 211, 22-35.

John, P., Brooks, B., Wand, C., & Schriever, U. (2013). Information density in bridge team communication and miscommunication - a quantitative approach to evaluate maritime communication. *WMU Journal of Maritime Affairs*, 122, 229-244.

Kilgarriff, A. (2005). Language is never, ever, ever, random. *Corpus linguistics and linguistic theory*, 12, 263-276.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different?. *Applied linguistics, 192*, 255-271.McEnery, T., & Wilson, A. (2001). Corpus linguistics: An introduction. Edinburgh: Edinburgh University Press.

Möckel S., Brenker M., Strohschneider S. (2014). Enhancing Safety through Generic Competencies. In *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation*, 81, 97-102

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. *Handbook of reading research*, 3, 269-284.

Noble, A., Vangehuchten, L., & Van Parys, W. (2011). Communication for maritime purposes: some exploratory results of a survey-based study on intercultural and linguistic features. *International Journal of Applied Linguistics*, 162, 111-133.

Potter, R. E., & Balthazard, p. A. (2002). Understanding human interaction and performance in the virtual team. *Journal of Information Technology Theory and Application JITTA, 41*, 6-30.

Pritchard, B. (2003). Maritime English syllabus for the modern seafarer: Safety-related or comprehensive courses? *WMU Journal of Maritime Affairs*, 2(2), 149-166.

Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. *Proceedings of the workshop on Comparing Corpora*. Stroudsburg: Association for Computational Linguistics, 1-6.

Seidlhofer, B., Breiteneder A., Klimpfinger, T., Majewski, S., Osimk-Teasdale, R., Pitzl, M. L., Radeka, M. (2013). The Vienna-Oxford International Corpus of English version

2.0 XML. Internet resource available at http://www.univie.ac.at/voice/, viewed on 30 June 2014.

Trenkner, p. (1996). IMO-Standard Marine Communication Phrases (SMCP) – an attempt to meet increased communication requirements of ship's officers. *Ninth International Maritime Lecturers' Association International Conference on Maritime Education and Training (IMLA 9)*, Kobe: International Maritime Lecturers' Association.

Trenkner, P., Cole, C. (2010). Raising the Maritime English bar: The STCW Manila Amendments and their impact on Maritime English. *Proceedings of the 22nd International Maritime English Conference*, Alexandria: International Maritime Lecturers' Association, 3-16.

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, 151, 8-25.

Weeks, F. (1997). Whither Maritime English. Ninth international IMLA Workshop on Maritime English (WOME 9), Malmö: International Maritime Lecturers' Association.

Linguistic measurement of cognitive load in maritime team communication by native and non-native speakers of English[1]

Peter John

Faculty of Maritime Studies, Jade University of Applied Sciences, Germany

Australian Maritime College, University of Tasmania, Australia

Benjamin Brooks

Australian Maritime College, University of Tasmania, Australia

Ulf Schriever

Australian Maritime College, University of Tasmania, Australia

Abstract

The research paper presents an assessment of individual cognitive load levels in simulated routine work tasks performed by native speakers and non-native speakers of English. This cross-sectional research is based on observational studies including audio-recorded nautical scenarios in full-mission bridge simulation. Primary task load experienced by simulation participants was measured by linguistic markers. Statistical hypothesis testing was conducted on the basis of quantitative content analyses and significant effects were expressed as the respective Probability of Superiority sizes. Findings show an increased cognitive load in non-native speakers while carrying out similar tasks in an almost identical environment.

key words: cognitive load, maritime simulation, bridge team communication, psycholinguistics, quantitative content analysis

1. Introduction

In the modern work environment a series of factors are challenging our ability to work efficiently and effectively. Many work environments are characterised by a

1 This is the author's original manuscript as submitted to the journal of Applied Ergonomics.

steadily increasing flow of information, often facilitated by technology with varying degrees of usability. In safety-critical workplaces levels of reliability must be high, with individuals and work teams needing to build and maintain situational awareness and support their decision-making (Chauvin et al., 2013; Visentini & Snidaro, 2013). At the same time, modern business processes can be said to develop in an increasingly international setting where English is used as the *Lingua Franca* for multi-national communication (Crystal, 2012).

These characteristics of the work environment can be hypothesised to have a marked impact on the 'cognitive load' that individuals and teams experience. Cognitive (work)load has been described by Hart & Staveland (1988, p140) as a "hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance". Needing to communicate in a non-native language simultaneously may further influence cognitive load, and compromise human performance. When the margins for error are small, this compromise can possibly have serious consequences.

Nowhere are the tendencies towards an increased information processing and *Lingua Franca* communication more visible than in the shipping industry. Today's mega ships are operated by small crews of some twenty members or fewer who work in shifts twenty-four hours a day, seven days a week. Technological advances have provided ships with multiple navigational aids which aim to support the nautical officers in making the right decisions. The shipping industry is also at the forefront of international work environments as nearly all crews of merchant ships are now multi-ethnic and multi-lingual (Noble et al., 2011a, 2011b). For this reason, the use of English as the *Lingua Franca* on board was made compulsory by the International Maritime Organisation (IMO) for discursive situations where crew members do not share a common language (International Maritime Organization, 2001, 2010).

To some extent, the shipping industry's highly technological and multi-national work conditions anticipate tendencies which have also become increasingly conspicuous in other work environments. Therefore, findings on cognitive load levels in a maritime setting may well be reproducible in other high-risk areas with a similar degree of internationalisation, such as aviation, military operations, or operating theatres.

2. Definition and measurement of cognitive load

Cognitive load (CL) is closely related to mental workload, a "multidimensional construct representing the load that performing a particular task imposes on the cognitive system" (Paas & van Merriënboer, 1994, p420). Both mental workload and cognitive load have been studied by a number of researchers using a variety of different methodologies. However, in spite of the numerous efforts to track mental workload and cognitive load, "there is no clearly defined, universally accepted definition" of these latent psychological variables, as Cain's (2007) review of this literature has clearly indicated. The ongoing debate is highlighted by Schmeck et alia (2014, p95) by saying that "[t]he question of how to assess cognitive load most validly and reliably is still leading to heated discussions among researchers".

While mental workload aims to describe a person's efforts to carry out a task, cognitive load embraces task performance plus the processing of new information, most often in a learning environment. Accordingly, Khawaja et alia make reference to Chandler & Sweller (1991) and define cognitive load as "the amount of mental demand imposed on a person by a particular task [...] associated with the limited capacity of the person's working memory *and the ability to process novel information*". In an instructional setting, Cognitive Load Theory (CLT) strives to explain "the learning of complex cognitive tasks, where learners are often overwhelmed by the number of information elements and their interactions that need to be processed simultaneously before meaningful learning can commence" (Paas et al., 2004, p1).

Given that cognitive load is not directly measurable, different techniques have been developed to capture this fugacious phenomenon by means of proxy variables. These techniques can broadly be categorised into three approaches which include subjective estimations (e.g. questionnaires, interviews), physiological effects (e.g. brain activity, pupil dilatation, sweat production, skin conductivity) and performance measures. The latter can be divided into the two sub-categories of primary task measures (e.g. error counting, completion time) and secondary task measures. Here, the term "secondary" refers to tasks subjects undertake while simultaneously performing their main or primary task (e.g. concurrent mental calculation). For an

overview on measuring techniques for cognitive load, see Paas et alia (2003) and Plass et alia (2010).

The maritime community has long identified the human element as a major contributory factor to promote safe navigation in a socio-technical environment (Grech et al., 2008). Cognition is believed to be at the core of decision-making processes which has lead to vivid research activities on cognitive processes and mental workload in the maritime domain (Sanquist et al., 1994; Lee & Sanquist, 2000; Itoh, 2001; Robert et al., 2002; Bjørkli, 2007; Kim & Hong, 2010; Uitterhoeve, 2011; Lenk et al., 2012; Tac et al., 2013; Stanton, 2014; Afeltowicz & Wachowski, 2015; Orlandi et al., 2015; Peysakhovich, 2015). Studies focussing on cognitive load have been conducted in the area of maritime education and training (MET; Grootjen et al., 2006, 2007), by analysing maritime accident reports (Embrey et al., 2006), and in real-life operations at sea (Bjørneseth et al., 2012, 2014). Neerincx et alia (2009) combine computer lab and sea-based data for calculating probability relationships between cognitive task load and performance.

3. Measuring cognitive load by linguistic variables

Most of the maritime research mentioned above makes inferences on cognitive load by means of proxy variables captured in primary tasks. The maritime area is especially apt for this type of data collection as all navigation is undertaken in a socio-technical environment and involves handling a series of electronic devices. The use of these devices can be logged and subsequently analysed and compared to task achievements which are generally clear and unambiguous (e.g. leaving a berth, entering a fairway, avoiding a collision with crossing traffic, etc.).

This research however follows a relatively recent approach to measure latent psychological variables linguistically. The development of the Linguistic Inquiry and Word Count (LIWC) software has increased the popularity of Quantitative Content Analysis (QCA) for scrutinising cognitive constructs through the use of discursive communication, with over 5,000 research papers published since the year 2000. QCA has been characterised as a valid and reliable technique to identify and compute frequencies of different word categories or grammar classes (Kracauer, 1952; Frey et al., 2000, pp236-241; Rourke & Anderson, 2004; Marsh & White, 2006). It converts

quantitative observations in language into metric variables which can be analysed statistically.

In Wickens' 4D multiple resource model (Wickens, 2008) QCA fits into the *auditory* modality with the *verbal* code on the perception side and into the *vocal verbal* code on the responding side.

4. Aims of the research and hypotheses

To date, no research has been reported which takes on the challenge of inferring cognitive load levels in a maritime setting based on QCA. The authors seek to fill this gap by comparing linguistic proxy variables of cognitive load (CL) observed for native speakers (NS) and non-native speakers (NNS) of English who are engaged in identical collaborative tasks. For this purpose, an approach presented by Khawaja et alia (2012) is adopted, who base their research on findings on cognitive models and their effect on language production and linguistic patterns as evidenced by Atkinson & Shiffrin (1968), Baddeley (1992, 2000, 2003), Kintsch, Sweller et alia (1998) and Patel, & Ericsson (1999).

In their study on Australian bushfire fighters, Khawaja et alia compare word count figures as well as agreement and disagreement word and pronoun frequencies in low load and high load tasks and report significant differences between the two task settings. They conclude that their study “provides encouraging evidence and presents some novel linguistic and grammatical features extracted from natural speech as potential indices of users’ experienced cognitive load” (Khawaja et al., 2012, p527).

Khawaja et alia look at the variation of manifest proxy variables within one group of participants while they carry out tasks of varying cognitive load levels. The present research adopts a different experimental design, studying two different groups, namely native and non-native speakers, while they perform identical routine tasks. This means that in contrast to the study cited above, the present research does not focus on different task difficulties but it analyses the cognitive load levels the participating teams experience while performing almost identical tasks. The complex socio-technical environment in which the tasks are performed will have an impact on participants regardless of their mother tongue, but the question if and to what extent non-native speakers of English experience a higher cognitive load level than native speakers is still

unknown. As the participants' native language constitutes the most significant difference between both groups, it is used as the independent variable. Metric linguistic variables which have been identified to manifest cognitive load levels (Pennebaker et al., 2003; Tausczik & Pennebaker, 2010) are analysed as dependent variables.

In line with Khawaja et alia's approach, the dependent variables correlated to cognitive load are formulated as null hypotheses as follows:

$H_0 1$: $WC_{NS}=WC_{NNS}$ where WC_{NS} is the word count of native speakers, and WC_{NNS} is the word count of non-native speakers.

$H_0 2$: $AW_{NS}=AW_{NNS}$ where AW_{NS} is the number of words expressing agreement by native speakers, and AW_{NNS} by non-native speakers.

$H_0 3$: $DW_{NS}=DW_{NNS}$ where DW_{NS} is the number of words expressing disagreement by native speakers, and DW_{NNS} by non-native speakers.

$H_0 4$: $SP_{NS}=SP_{NNS}$ where SP_{NS} is the number of singular pronouns by native and SP_{NNS} by non-native speakers

$H_0 5$: $PP_{NS}=PP_{NNS}$ where PP_{NS} is the number of plural pronouns by native and PP_{NNS} by non-native speakers

Additionally, the LIWC category of “cognitive words” has been added (as used by Sharp, 2004; Boals et al., 2011):

$H_0 6$: $CW_{NS}=CW_{NNS}$ where CW_{NS} is the number of cognitive words by native and CW_{NNS} by non-native speakers

5. Data collection

This research is based on empirical speech data collected by audio-recording trainee nautical officers during their normal training sessions in full-mission bridge team simulation. This simulation environment truly replicates a merchant ship's bridge and enables participants to engage in a meaningful, constructivist learning process. Full-

mission simulation is an integral part of Maritime Education and Training (MET). It provides an optimal environment for researching social interaction because of its possibility to create a homogeneous primary task workload while not imposing any secondary workload on participants (Brünken et al., 2002; van Gog et al., 2011). Hence, the controlled environment of a research experiment is obtained while allowing for observational studies without introducing possible biases caused by the presence of the research team. Audio-recording teams in full-mission simulation is a normal procedure in MET as speech data are often used for debriefing the exercises to spot mistakes and suggest improvements. Further, most ships record voice and data communications through the use of the Voyage Data Recorder (VDR), so the research environment is simply a representation of the actual work environment in this regard.

Participants were recorded at the Faculty of Maritime Studies of Jade University of Applied Sciences in Germany and at the National Maritime College of Ireland. At each institution, a total of 20 volunteers participated in the exercises. All participants in Ireland were native speakers of English whereas all participants in Germany were non-native speakers of English.

The data were collected with the informed consent of all participants involved and in compliance with the Ethics stipulations of both institutions and with the Social Sciences Human Research Ethics regulations of the University of Tasmania (Australia). No participants withdrew their consent during or after the recordings thus eliminating any attrition effects.

The audio recordings were transcribed by the corresponding author and validated by his co-authors. For this purpose, random transcript excerpts were compared with the original audio recordings. Ambiguous or unintelligible words were marked with a wildcard character.

As two different groups of individuals are compared, potential biases caused by a possibly heterogeneous group composition need to be considered. It is understood that the smaller the difference in the composition of the two groups, the fewer biases are introduced into the research. For this purpose, meta-data were surveyed from participants which included their gender, age, sea experience and their last work on board a ship. Participants were also asked about the crew composition of their ships,

namely if the crew was multinational, and if English was the prevalent language spoken amongst the nautical officers on board.

6. Data analysis

The collected verbal discourse data is analysed in two steps. Firstly, control variables are examined which may have biasing effects on the conducted comparison and secondly, the identified proxy variables for cognitive load are compared. The control variables include demographic meta-data of the participants and their linguistic profile.

6.1 Demographic meta-data and work experience

6.1.1 gender and age

All participants were male with a median age of 25 years in Germany (m=26, sd=4) and 26 years in Ireland (m=31, sd=8). A Mann-Whitney-U analysis of variance[2] lead to a non-significant result of p=0.13 so that the age distribution can be assumed to originate from the same population. Hence, this control variable is considered sufficiently homogeneous for a between-groups comparison without the need to consider it an influencing factor.

6.1.2 recent and total sea experience

All participants were asked to state the year they had last worked on board a ship. The German participants gave 2012 as the median year (m=2012, sd=1) while the future Irish officers stated 2013 as the median year (m=2013, sd=1). The corresponding hypothesis testing lead to p=0.40, so that this variable does not need to be controlled either.

The total time participants had worked on board a ship prior to the recorded exercises is another variable that needs to be controlled as this exposure to real life work conditions can possibly have a decisive influence on the way participants communicate and thus on the analysed linguistic variables. Here, the meta-data revealed a clear difference with a median sea experience of 13 months (m=14, sd=2) for the German participants and 40 months (m=108, sd=142) in the case of the Irish.

2 The non-parametric Mann-Whitney-U analysis of variance was used for all hypothesis testing as no pair of variable sets was neither found to be normally distributed (Anderson-Darling test, α=0.1) nor homoscedastic (Levene test, α=0.1).

As expected these strongly diverging figures lead to a significant difference, with $p<0.00$ so that the null hypothesis of no difference between the two groups must be rejected.

A multiple regression analysis was carried out to determine whether the much longer exposure of the Irish participants to a real-life work experience could possibly introduce bias to the analysis of cognitive load levels. The regression analysis computed the correlation of the total sea experience of participants with their individual frequencies of personal pronouns together and separately, as well as with their frequencies of cognitive words, and with their agreement and disagreement word frequencies. The multiple regression analysis lead to a coefficient of multiple determination of $R^2=0.26$, with an adjusted $R_a^2=-0.03$. The corresponding analysis of variance resulted in a non-significant value of $p=0.59$ for the cognitive load proxies together, and in non-significant values of between $p=0.18$ and $p=0.76$ (m=0.38, sd=0.26) for the proxies individually.

This finding is rather interesting because it shows that the frequencies of words expressing cognitive load are hardly related to the time the participants were exposed to their real work environment at sea. In other words, even though the median sea time of the Irish participants was more than double that of the German participants, this much longer exposure very marginally influenced the cognitive load levels they experienced during the full-mission simulation exercises. A possible explanation for this effect could be a linguistic saturation that had been reached by the participants, as they had already acquired a sufficient degree of nautical idiomaticity before finishing their sea experience. Given the weak correlation, a possible impact of the total sea time on the cognitive load levels was deemed of little importance and thus not further analysed.

6.1.3 multinational composition of ship crew and language spoken amongst officers

Apart from the individual time and year of their last work experience, participants also rated the time they had sailed with multinational crews on a five-point Likert scale[3]. The exposure to a multinational work environment resulted in a median of one (“always”) both for the German (m=1.5, sd=0.8) and the Irish participants (m=1.6,

3 A five-point Likert scale was used stating the following options: 1=always, 2=mostly, 3=half of the time, 4=hardly ever, 5=never.

sd=0.8). Not surprisingly, the corresponding null hypothesis of no difference in the multinational exposure was accepted with p=0.46.

When asked to state the language which was mostly spoken amongst the officers on board, both participant groups also answered with a median of one ("always"[4]; German group m=1.2, sd=0.7; Irish group m=1.0, sd=0), leading to an acceptance of the null hypothesis of no difference with p=0.48.

The analysis of the collected meta-data reveals that the Irish and the German participants formed two quite similar groups in terms of demographic data and work background. The only statistically significant difference in their total work experience appears to be very weakly correlated to the groups' cognitive load proxies so that a possible influence on the latter was disregarded.

6.2 Linguistic profile

Apart from analysing the participants' meta-data, a linguistic profile was determined for each group. These two linguistic profiles reflect the idiosyncratic communication patterns of each team by including the teams' mean word production time, their word production per utterance, type-token ratio, lexical density, key word ratio and their special part-of-speech diversity index, following an approach outlined by John et alia (2017).

Both the meta-data and the linguistic profiles are used as control variables to determine their correlation and hence a possible influence on the proxy variables for cognitive load levels. The meta-data analysis has shown that both groups present similar characteristics as to the personal and work background of the audio-recorded participants. In the following, their linguistic behaviour is studied to determine whether it is possibly correlated to their cognitive load proxies. For this reason, some linguistic key variables are compared statistically by means of Mann-Whitney-U hypothesis testing following Anderson-Darling tests for normality and Levene tests for homoscedasticity (α=0.05 was assumed for both test types) . Between-groups effect sizes were computed as the Probability of Superiority (PS) which does neither

4 The five-point Likert scale included the following answers: 1=English, 2=my mother tongue (if not English), 3=another language I speak, 4=another language I do not speak, 5=no answer

assume a normal nor a homoscedastic distribution of the speech data (cf. Grissom & Kim, 2005). Table 1 summarises the statistical testing of the linguistic variables.

Linguistic feature	Non-native speakers (NNS)	Native speakers (NS)	Effect size
Median word production time in seconds	1.19	1.77	NNS<NS
Word production per utterance*[2, 4]	median=5 m=6.45, sd=5.65	median=6 m=7.46, sd=7.08	PS=0.46
Type-token ratio*[1, 4]	0.042	0.067	PS=0.25
Lexical density*[2, 4]	median=0.41 m=0.47, sd=0.31	median=0.40 m=0.44, sd=0.30	PS=0.53
Key word density*[2, 3]	median=0.21 m=0.27, sd=0.24	median=0.21 m=0.25, sd=0.34	PS=0.54
Special part-of-speech diversity index**[1, 4]	median=1.00 m=0.92, sd=0.20	median=1.00 m=0.91, sd=0.24	PS=0.51

*(*p<0.00; **p=0.01; [1]normally distributed, [2]not normally distributed; [3]homoscedastic; [4]heteroscedastic)*

table 1: Linguistic differences between both participant groups

6.2.1 word production and type-token ratio

During the total recording time of ten hours the German teams produced a total of 43,019 words which equates to 1.19 words per second or 72 words per minute. The word production of the Irish teams was much higher, with 63,871 words in total, i.e. 1.77 words per second or 106 words per minute. No hypothesis testing was conducted at this stage as these figures do not reflect any distribution. At an utterance level, the hypothesis testing resulted in a significant difference with $p<0.00$. However, the computed effect size of PS=0.46 indicates that in 46 out of a hundred cases the German teams produced fewer words per utterance than the Irish teams. This value is very close to 0.50 which would mean that the probability for a longer utterance to

occur in either group is identical, so that no statistical difference is observed. Given the small effect size a possible impact of this linguistic variable was considered negligible.

The type-token ratio (TTR) expresses the ratio of different words ("types") to the total word count ("tokens"). As the type-token ratio is known to be influenced by the total text length and given that the Irish text sample is much bigger than the German sample, the former was reduced to the latter's size by random-sampling ten fragments with replacement. The resulting mean PS effect size of 0.25 corresponds to a very significant difference of 1843 word types for the German and 2899 word types for the Irish group. This finding does not really come as a surprise as native speakers can be expected to use a richer or more varied vocabulary than non-native speakers who are more likely to apply the coded language reflected in the Standard Marine Communication Phrases (International Maritime Organization, 2001).

To eliminate doubts as to whether a higher number of word types automatically results in an increased frequency of cognitive load proxies, a multiple regression analysis was conducted on the individual vocabulary growth figures and on the cognitive load proxies stated above. The coefficient of multiple determination was R^2=0.36 and the adjusted coefficient R_a^2=0.15. A subsequent analysis of variance lead to p=0.13, with probability values ranging from p=0.03 to p=0.92 (m=0.66, sd= 0.29) for the individual cognitive load proxy variables. Given the very low R_a^2 value and the acceptance of the null hypothesis it was decided not to investigate a possible influence on cognitive load further.

6.2.2 lexical and key word frequency

Lexical density is the ratio of content words over the total word count. This ratio is useful to estimate the extent to which information is exchanged (John el al., 2013) while the key word density represents the degree of idiomaticity in a discourse situation (ibid.). For both linguistic variables the calculated PS effect size was quite balanced, with a value of 0.53 and 0.54, respectively. This close proximity differs very clearly from effect sizes identified in previous research between maritime full-mission simulation and verbal discourse situations outside a nautical setting for which John et alia (2017) report PS effect sizes of 0.70 (lexical density) and 0.68 (key word density).

For this reason, the small variations in the distribution of lexical density and key word ratios were not studied further for a possible correlation with cognitive load figures.

6.2.3 part-of-speech diversity

While the linguistic features analysed above all relate to vocabulary items, part-of-speech diversity focuses on the number of different word classes per utterance as a means to quantify the utterances' grammar structure. Comparing the German with the Irish group by means of the special part-of-speech diversity index (John & Brooks, 2014) resulted in a nearly identical distribution with an effect size of PS=0.51. Again, no further analysis was conducted on a possible correlation.

Khawaja et alia's research measured cognitive load levels linguistically while the same subjects carried out tasks of varying difficulty. In contrast, this research studies two different groups of participants which have been found to differ only marginally from each other in terms of their demographic and work background as well as regarding their linguistic behaviour. The latter, somewhat surprising finding can possibly be explained by the good English proficiency of the non-native speakers who prior to the recorded exercises all had had more than ten years of English language learning at school and at the university plus a median work experience of 13 months in a multinational, multi-ethnic work environment where English was used extensively. On the other hand, it highlights the robustness of quantitative content analysis as a methodology for extracting frequencies which can be analysed statistically as metric variables. In previous research by John et alia (2017) very high PS effect sizes were reported between linguistic profiles of maritime and non-maritime communication. The present findings corroborate the proximity of language patterns in a specific ESP discourse situation which contrasts starkly with general everyday English.

6.3 Comparison of cognitive load levels

The statistical analysis of the demographic meta-data and the participants' linguistic profiles has shown that their influence on the proxy variables of cognitive load can be disregarded. In the following, the question is studied whether these demographically very similar groups also experience a similar level of cognitive load while they perform identical routine tasks. If non-native speakers do experience a higher cognitive load the proxy variables will lead to significant between-groups differences and high effect sizes.

On the contrary, if both groups experience similar cognitive load levels, the effect sizes will centre around PS values of 0.50.

In demanding tasks which impose a high cognitive load, participants can be hypothesised to communicate more with each other in order to clarify challenging situations and discuss possible solutions, thus achieving a common situational awareness. In their research on bushfire management tasks, Khawaja et alia report a significant difference of 23% more words per sentence (i.e. utterance) in high load tasks. However, in a direct comparison between native and non-native speakers of English in a maritime setting it is the native speakers who are found to produce 33% more words in total (see table 2).

Linguistic features	**Native speakers**	**Non-native speakers**	**Difference**	
Word count[1]	63,871	43,019	33%	
Words per utterance[2]	6.45	7.46	-16%	PS=0.55*
Cognitive words[3]	11.62	13.93	20%	PS=0.31*
Agreement words[3]	4.93	6.25	-27%	PS=0.38*
Disagreement words[3]	1.24	1.44	-16%	PS=0.48*
First-person singular pronouns[3]	1.60	2.92	-83%	PS=0.26*
First-person plural pronouns[3]	2.76	3.52	-28%	PS=0.31*
Total effect size:				PS=0.33

[1]In number of words, [2]median, [3]in % of word count, *$p<0.000$

table 2: Summary of linguistic features of cognitive load (n=40)

However, a comparison of the number of words per utterance reveals that non-native speakers produced significantly longer utterances than non-native speakers. Hence, the null hypothesis $H_0 1$: $WC_{NS}=WC_{NNS}$ must be rejected, albeit with a PS effect size of only 0.55 (16% difference).

Looking at agreement and disagreement words, the group experiencing higher cognitive load is expected to produce fewer of the former and more of the latter, in line with findings by Khawaja et alia. Here, both null hypotheses ($H_0$2: AW_{NS}=AW_{NNS}, $H_0$3: DW_{NS}=DW_{NNS}) must be rejected with effect sizes of PS=0.38 (27% difference) and PS=0.48 (16% difference), respectively. Non-native speakers clearly produce more agreement words but they also produce slightly more disagreement words.

The use of plural personal pronouns expressing group identity (*we, us, our*) have been found to increase in high load situations while the frequency of singular personal pronouns (*I, me, mine*) can be expected to decrease. In this research, the group experiencing higher cognitive load levels can therefore be expected to present a higher frequency of plural pronouns and a lower frequency of singular pronouns. Again, significant differences are found for both hypotheses ($H_0$4: SP_{NS}=SP_{NNS}, $H_0$5: PP_{NS}=PP_{NNS}). The non-native speaker groups produced more plural pronouns (PS=0.31, 28% difference) but they also produced a higher frequency of singular pronouns (PS=0.26, 83%). In contrast to agreement and disagreement word frequencies, a similar effect size is observed in singular and plural personal pronouns.

Lastly, a comparison of the LIWC category of "cognitive words" leads to a significant difference between both groups with a 20% higher value for the non-native speakers ($H_0$6: CW_{NS}=CW_{NNS}). Again, the PS effect size of 0.31 reveals a marked difference in the frequency of proxy variables for cognitive load, with non-native speakers experiencing a clearly increased strain to process novel information.

Table 2 summarises the statistical findings of the between-groups comparisons.

7. Findings and discussion

This research strives to answer the question to what extent two groups of work teams experience different cognitive load levels while they carry out routine tasks in an identical environment if the clearest difference between both groups is their native language. For this reason, two groups are compared which do neither present significant differences in their demography and work experience nor in their ESP (here: Maritime English) idiomaticity levels.

Communication is considered one – if not the most – important element of social interaction in team work. By following a research method adopted in a study on safety-critical collaboration in bush fire management, linguistic proxy variables for cognitive load are studied and significant findings are found for all analysed variables. Cognitive word frequencies according to the respective LIWC category are significantly higher in the non-native speaker group than in the native speakers which hints at a clearly increased cognitive load level because of their non-nativeness alone. The findings are somewhat less clear for the use of agreement and disagreement words where the non-native speaker group produces higher frequencies in both categories. The same applies to the use of personal pronouns where the frequencies are also higher in both non-native speakers' categories. Although these findings do not replicate the findings by Khawaja et alia, the authors argue that the frequencies of these word categories express a higher cognitive load in themselves. The non-native teams produced fewer words than the native-speakers and nevertheless they expressed more agreement and disagreement, and they referred more often to themselves in singular and in plural. These increased frequency levels are clear indicators of cognitive processes involved in collaborative decision-marking and developing shared mental models as reflected by the communication transcript excerpt displayed in table 3.

(4720) shipmaster	So, we must make a new position.
(4721) officer	We get four miles, yes?
(4722) shipmaster	South Anchorage, so.
(4723) officer	It would be better to tell them now, just give them a rough position, that they are steering here.
(4724) shipmaster	From this buoy.
(4725) officer	Where, where is the South Anchorage?
(4726) shipmaster	Is that here?
(4727) officer	No, he said on this, this chart number three .
(4728) shipmaster	But, number three is not, here is number three.
(4729) officer	No, it's chart number three.
(4730) helmsman	* this chart.

(4731) officer	Now we need to find this.
(4732) helmsman	Shall I change the course?
(4733) officer	We need just to find the South Anchorage first.
(4734) shipmaster	I understood the South Anchorage's here.
(4735) officer	No, no, we have a better chart.
(4736) officer	South, South Anchorage off Heligoland.
(4737) officer	Where is it?
(4738) officer	Here is one, Anchorage.
(4739) shipmaster	Elbe approach.
(4740) officer	There is Elbe.
(4741) shipmaster	Yeah.
(4742) officer	Here, South.
(4743) shipmaster	Uh, okay.
(4744) shipmaster	It's this position, yeah?
(4745) officer	South Anchorage was leaving *.
(4746) shipmaster	Erm, yeah.

table 3: communication excerpt[5]

The mean PS effect size across all linguistic variables of PS=0.33 has quantified the increased cognitive load levels for non-native speaker teams in an almost identical workplace and task setting. This confirms that non-native speakers experience a significantly higher cognitive load when carrying out the same tasks than teams who communicate in their native language. The finding might not be unexpected from a qualitative perspective, but by quantifying metric variables the magnitude of the observed effects can be compared. Sometimes, this leads to unexpected results. The comparison of the participants' meta-data has shown that two apparently different groups of teams may produce non-significant differences in their control variables as in the case of the participants' on-board language exposure on the dependent linguistic variables. This finding was not anticipated by the authors who have several decades of

5 First person pronouns and agreement/disagreement words are underlined.

experience in teaching future nautical officers and who did not expect students to reach a linguistic saturation point so soon.

In a safety-critical environment like the shipping industry, higher cognitive load levels can be decisive for success or failure. As in many other safety-critical areas, in the shipping industry failure can cause the loss of lives and substantial damage to the environment. Inferring cognitive load levels by means of measurable linguistic features and quantitative content analysis methods indeed provide "encouraging evidence" and a robust and reliable way of quantifying latent psychological proxy variables.

Future research might be able to correlate linguistic proxy variables to primary task measures to triangulate research methods and obtain a degree of calibration of cognitive load measurements.

8. References

Afeltowicz, Ł., & Wachowski, W. (2015). How Far we Can Go Without Looking Under the Skin: The Bounds of Cognitive Science. *Studies in Logic, Grammar and Rhetoric*, *40*(1), 91-109.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The psychology of learning and motivation*, *2*, 89-195.

Baayen, R. H. (2001). Word frequency distributions Vol. 18. Springer.

Baddeley, A. (1992). Working memory. *Science*, 255 (5044), 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, 4(11), 417-423.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews neuroscience*, 4(10), 829-839.

Bjørkli, C. A. (2007). Joint Cognitive Systems at Sea: A Study of High Speed Craft Operation. Doctoral thesis at the Norwegian University of Science and Technology.

Bjørneseth, F. B., Renganayagalu, S. K., Dunlop, M. D., Homecker, E., & Komandur, S. (2012, September). Towards an experimental design framework for evaluation of

dynamic workload and situational awareness in safety critical maritime settings. *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, 309-314.

Bjørneseth, F. B., Clarke, L., Dunlop, M. D., & Komandur, S. (2014). Towards an understanding of operator focus using eye-tracking in safety-critical maritime settings. *International Conference on Human Factors in Ship Design & Operation*

Boals, A., Banks, J. B., Hathaway, L. M., & Schuettler, D. (2011). Coping with stressful events: Use of cognitive words in stressful narratives and the meaning-making process. *Journal of Social and Clinical Psychology*, 30(4), 378-403.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental psychology*, *49*(2), 109-119.

Cain, B. (2007). A review of the mental workload literature. Defence Research and Development. Toronto.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and instruction*, 8(4), 293-332.

Chauvin, C., Lardjane, S., Morel, G., Clostermann, J. P., & Langard, B. (2013). Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. Accident Analysis & Prevention, 59, 26-37.

Crystal, D. (2012). English as a global language. Cambridge University Press.

Embrey, D., Blackett, C., Marsden, P., & Peachey, J. (2006). Development of a human cognitive workload assessment tool. Dalton (UK): Human Reliability Associates.

Frey, L., Botan, C. H., & Kreps, G. (2000). Investigating communication. *NY: Allyn & Bacon*.

Grech, M., Horberry, T., & Koester, T. (2008). *Human factors in the maritime domain*. CRC Press.

Grissom, R. J., & Kim, J. J. (2005). Effect sizes for research. A broad practical approach. Mah.

Grootjen, M., Bierman, E. p. B., & Neerincx, M. A. (2006). Optimizing cognitive task load in naval ship control centres: Design of an adaptive interface. *IEA 16th World Congress on Ergonomics*.

Grootjen, M., Neerincx, M. A., & Marckelbach, M. (2007, October). Dynamic task load scheduling for platform control and navigation on a naval ship. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 51(21), 1463-1467).

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A.H.N. Meshkati (Ed.), Human mental workload (139–183). Oxford, UK: North-Holland.

International Maritime Organization (2001). Resolution A.918(22): IMO Standard Marine Communication Phrases, London.

International Maritime Organization (2010). International Convention on Standards of Training, Certification, and Watchkeeping for Seafarers, 1978, Manila Amendments 2010, London.

Itoh, K., Yamaguchi, T., Hansen, J. P., & Nielsen, F. R. (2001). Risk analysis of ship navigation by use of cognitive simulation. *Cognition, Technology & Work*, *3*(1), 4-21.

John, P., Brooks, B., Wand, C., & Schriever, U. (2013). Information density in bridge team communication and miscommunication – a quantitative approach to evaluate maritime communication. *WMU Journal of Maritime Affairs*, 12, 229–244.

John, P., & Brooks, B. (2014). Lingua Franca and its Grammar Footprint: Introducing an Index for Quantifying Grammatical Diversity in Written and Spoken Language. *Journal of Quantitative Linguistics*, 21(1), 22-35.

John, P., Brooks, B., Schriever, U. (2017). Profiling maritime communication by non-native speakers: A quantitative comparison between the baseline and standard marine communication phraseology. *English for Specific Purposes,* 47, 1–14.

Khawaja, M. A., Chen, F., & Marcus, N. (2012). Analysis of collaborative communication for linguistic cues of cognitive load. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54(4), 518-529.

Kim, H., & Hong, S. (2010). Collision Scenario-based Cognitive Performance

Assessment for Marine Officers. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 4(1), 73-77

Kracauer, S. (1952). The challenge of qualitative content analysis. *Public opinion quarterly*, 631-642.

Lee, J. D., & Sanquist, T. F. (2000). Augmenting the operator function model with cognitive operations: Assessing the cognitive demands of technological innovation in ship navigation. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *30*(3), 273-285.

Lenk, J. C., Droste, R., Sobiech, C., Lüdtke, A., & Hahn, A. (2012). Towards cooperative cognitive models in multi-agent systems. *International Conference on Advanced Cognitive Technologies and Applications*, 67-70.

Marsh, E. E., & White, M. D. (2006). Content analysis: A flexible methodology. *Library trends*, *55*(1), 22-45.

Neerincx, M. A., Kennedie, S., Grootjen, M., & Grootjen, F. (2009). Modeling the cognitive task load and performance of naval operators. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience,* 260-270.

Noble, A., Vangehuchten, L., & van Parys, W. (2011a). Intercultural Competence and Effective Communication at Sea: An Invitation to Celebrate Diversity on Board. Proceeding of IMEC 23, 131.

Noble, A., Vangehuchten, L., & van Parys, W. (2011b). Communication for maritime purposes: some exploratory results of a survey-based study on intercultural and linguistic features. ITL. International journal of applied linguistics, (162), 111-133.

Orlandi, L., Brooks, B., & Bowles, M. A (2015). A Comparison of Marine Pilots' Planning and Manoeuvring Skills: Uncovering Mental Models to Assess Shiphandling and Explore Expertise. Journal of Navigation, 68(05), 897-914.

Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and motor skills*, *79*(1), 419-430.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. Educational psychologist, 38(1), 1-4.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. Instructional science, 32(1), 1-8.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. Annual review of psychology, 54(1), 547-577.

Peysakhovich, V., Vachon, F., Vallières, B. R., Dehais, F., & Tremblay, S. (2015). Pupil Dilation and Eye Movements Can Reveal Upcoming Choice in Dynamic Decision-Making. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 59(1), 210-214.

Plass, J. L., Moreno, R., & Brünken, R. (2010). Cognitive load theory. Cambridge University Press.

Robert, G., Hockey, J., Healey, A., Crawshaw, M., Wastell, D. G., & Sauer, J. (2003). Cognitive demands of collision avoidance in simulated ship control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(2), 252-265.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, *52*(1), 5-18.

Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2015). Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. Instructional Science, 43(1), 93-114.

Sanquist, T. F., Lee, J. D., & Rothblum, A. M. (1994). Cognitive analysis of navigation tasks: A tool for training assessment and equipment design. *Battelle Human Affairs Research Centres*, Seattle.

Sharp, W. G., & Hargrove, D. S. (2004). Emotional expression and modality: an analysis of affective arousal and linguistic output in a computer vs. paper paradigm. *Computers in human behavior*, 20(4), 461-475.

Stanton, N. A., & Bessell, K. (2014). How a submarine returns to periscope depth: Analysing complex socio-technical systems using Cognitive Work Analysis. *Applied ergonomics*, 45(1), 110-125.

Tac, U., Tavacioglu, L., Bolat, P., Kora, O. K., & Bolat, F. (2013). Monitoring seafarer's cognitive performance under stressor factors during a voyage by automated neuropsychological assessment metrics. *Universitatii Maritime Constanta. Analele*, 14(20), 291.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. Journal of language and social psychology, 29(1), 24-54.

Uitterhoeve, W., Croes-Schalken, M., & Ten Hove, D. (2011). Physiological Measurement Applied in Maritime Situations: A Newly Developed Method to Measure Workload on Board of Ships. *HCI International 2011–Posters' Extended Abstracts*, 327-331.

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied cognitive psychology*, *25*(4), 584-587.

Visentini, I., & Snidaro, L. (2013). New Trends for Enhancing Maritime Situational Awareness. *Prediction and Recognition of Piracy Efforts Using Collaborative Human-Centric Information Systems*, 193-200.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3), 449-455.

Speech acts in professional maritime discourse: A pragmatic risk analysis of bridge team communication directives and commissives in full-mission simulation[1]

Peter John

Faculty of Maritime Studies, Jade University of Applied Sciences, Germany

Australian Maritime College, University of Tasmania, Australia

Benjamin Brooks

Australian Maritime College, University of Tasmania, Australia

Ulf Schriever

Australian Maritime College, University of Tasmania, Australia

Abstract

The paper studies verbal maritime communication by categorising spontaneous professional discourse observed in co-operative full-mission simulation exercises into the illocutionary points of commissives and directives according to Searle's original classification. The research adopts a Corpus Pragmatics approach by combining vertical Corpus Linguistics methods with horizontal Pragmatics analyses. Between-group analyses of speech acts by native and non-native speakers of English are carried out and possible risks of miscommunication classified and compared. On the basis of the circular Osgood & Schramm communication model the sender-receiver interaction is investigated for either speaker group. Findings include both quantitative and qualitative between-group differences in locutionary, illocutionary and perlocutionary speech acts. These differences are evaluated as causal factors in effective communicative acts and as contributory factors for miscommunication in the maritime domain.

key words*: maritime team communication, professional discourse, Corpus Pragmatics, speech act theory*

1 The original article has been published by the Journal of Pragmatics under a Creative Commons licence: https://doi.org/10.1016/j.pragma.2018.11.013

1. Introduction

Team communication in the workplace plays a decisive role in human performance. Whilst professional team discourse fulfils a range of social functions (Holmes, 2005; Kraut et al., 1990; Li, 2000; Lynch, 2002) its main purpose is to contribute to a successful completion of work-related tasks. In this respect, the specific linguistic patterns employed by the discourse community of any professional domain can be expected to pursue the goal of effective communication as a means of task accomplishment (cf. Hoover, 2002).

In the safety-critical environment on board sea-going ships, effective team communication is of vital importance to safety as it prevents damage to vessels, injuries to their crews and environmental pollution (Bocanegra-Valle, 2011; de la Campa Portela, 2005; Jurkovič, 2015; Pritchard and Kalogjera, 2000). Conversely, miscommunication in the maritime domain has been identified as a major contributory factor to shipping accidents (cf. John et al., 2013; McCallum et al., 2000; McCrae, 2009; Pyne and Koester, 2005). For this reason, the International Maritime Organization (IMO), the United Nations body charged with maritime regulation, has made substantial efforts to reduce ambiguous language patterns by first developing the Standard Marine Navigational Vocabulary (International Maritime Organization, 1978) and later the Communication Phrases (SMCP, International Maritime Organization 2002). The SMCP were adopted by the IMO in the year 2001 for use by officers in charge of a watch on ships of 500 gross tonnage or more and were published in 2002. By simplifying and removing any ambiguities from natural language, the SMCP constitute a coded English language variety to be used by seafarers internationally (cf. Gustafsson, 2004; John et al., 2017; Noble, 2015).

Bridge team communication can be considered a sub-genre of Maritime English which relates to the verbal exchange of information by the navigating crew (on the bridge) of a ship[2]. It helps nautical officers to develop a shared understanding required to navigate a ship safely. A disruption of this information exchange due to an error in encoding or decoding a verbal message causes an incongruent mental representation of the

2 Bocanegra-Valle divides Maritime English into “five different subvarities according to the specific purpose they serve within the maritime context” (2013, p3580). Bridge Team Communication belongs to the subvariety of “[i]nternal (intra-ship or onboard) communication” (Bocanegra-Valle 2013, p3580).

navigational situation which potentially leads to erroneous and possibly dangerous decision-making processes (Balmat et al., 2011; Horck, 2004; Velasquez and Hester, 2013).

Verbal team communication can be visualised by means of the well-known Osgood & Schramm communication model (1954), whereby a message is transmitted successfully if it has been encoded appropriately by the transmitter and decoded correctly by the receiver. Following this model, a message (i.e. a communicative act) can be assumed to have been effective if encoded, transmitted and decoded correctly. Accordingly, communication can be deemed to have been ineffective if errors occur in any of its three phases of language production, language transmission or language reception (cf. Lent, 2013; McQuail and Windahl, 1993).

The Osgood & Schramm model of communication allows for a suitable framework to study speech acts in a maritime setting. The term “speech act” has been referred to by Davies (2005, p122-123) as the “smallest unit of analysis in conversational interaction”. Baker and Ellece (2011, p138) refer to speech acts as “utterances which perform various social functions such as requesting, greeting, advising, complaining, warning and so on”. Speech acts are an integral part of the speech act theory defined by Austin (1962) and developed further by Searle (1969). Davies (2005, p122-123) summarises the theory's underlying principle as follows: “[T]here are three types of speech act: the locutionary act (the basic literal meaning of an utterance), the illocutionary act (what the speaker intends by the utterance) and the perlocutionary act (the actual effect the utterance has on the hearer)”.

In theory, the coded English language variety promulgated by the IMO aims to remove any discrepancies between locutionary and illocutionary speech acts in bridge team communication. This means, that on the language production side of the Osgood & Schramm model speakers are expected to produce utterances without any ambiguities so that their locutionary language production and illocutionary meaning be identical. On the receiving end of the information exchange, nautical officers shall confirm perlocutionary acts by using a closed-loop communication whereby the receiver repeats the message uttered by the transmitter (hence closing the communicative loop). This affirmative communication method is commonplace in shipping (cf. Chauvin et al.,

2013; Chawla, 2015) and practised intensively in Bridge Resource Management courses which aim to improve the interaction of team members.[3]

The question arises as to what extent the theoretical construct outlined above can actually be observed in spontaneous speech acts by bridge team members. Preliminary research on bridge team communication introduced "a quantitative methodology to calculate and weight utterances for evaluating [the] information content" of maritime discourse based on lexical and key word densities (John et al., 2013, p242). The research approach was in line with the linguistics field of Semantics, by which locutionary acts were investigated in a literal manner, thus assuming no differences between locutionary and illocutionary speech acts.

This research sets out to study speech acts from the perspective of Pragmatics, which according to Paltridge (2006, p3) "is interested in what people mean by what they say, rather than what words in their most literal sense might mean by themselves". The research undertakes to identify possible miscommunication by singling out differences between locutionary and illocutionary speech acts. It also verifies if the "perlocutionary effect" of the corresponding speech acts corresponds to their "perlocutionary intention" (Bach, 1990, p397; Bach and Harnish, 1979). Communication patterns of two sociolinguistic groups, namely native and non-native speakers of English, are analysed and differences and similarities between these two groups are quantified.

2. Methodology

This paper investigates speech acts of naturalistic maritime communication recorded in full-mission simulation exercises. These exercises simulate the socio-technical work environment of a real ship's bridge. The analysis is carried out on a spoken corpus of bridge team discourse including native (L1) and non-native (L2) speakers of Maritime English. The spoken corpus contains verbatim transcripts of authentic, synchronic language use in standard navigation exercises. It consists of 43,019 word tokens produced by twenty German (L2) students of Nautical Sciences during a total recording time of 10h, and of 63,871 word tokens produced by twenty Irish (L1) students of Nautical Sciences during an identical total recording time. The special

3 At this point it is worth mentioning that the Standard Marine Communication Phrases (SMCP) have been paramount in introducing a communicative form which team members also adopt in situations not specifically covered by the SMCP.

spoken corpus is limited to standard team work tasks and has been studied in previous Corpus Linguistics research for lexical and key word densities, vocabulary growth and part-of-speech diversity (John et al., 2017). The empirical speech data was audio-recorded in the years 2013 and 2014 at Jade University of Applied Sciences in Germany and at the National Maritime College of Ireland (cf. John et al., 2016 for a detailed description of the data collection process and meta-data of exercise participants).

Corpus Linguistics traditionally pursues a "vertical-reading methodology" (Rühlemann and Aijmer, 2015, p8) on "authentic language data, stored in extensive computer corpora, as the basis for linguistic research" (Rühlemann and Aijmer, 2015, pi). This vertical approach in Corpus Linguistics enables researchers to establish frequencies of linguistic features which can subsequently be analysed by means of quantitative methods. To the contrary, Pragmatics most often adopts a "horizontal-reading methodology" Rühlemann and Aijmer, 2015, p3) in order to engage in "the art of the analysis of the unsaid" (Mey, 2001, p245). As Pragmatics research intends to infer the "meaning-in-context" (Bublitz and Norrick, 2011, p*v*) of utterances rather than studying their surface structure, research focuses on speakerelistener interaction which usually covers several utterances (cf. Bublitz and Norrick, 2011; Fasold, 1990; Leech, 1983; Levinson, 1983).

For the investigation of speech acts included in a spoken corpus on bridge team discourse a Corpus Pragmatics methodology has been applied. Corpus Pragmatics employs a vertical Corpus Linguistics approach to identify locutionary speech acts which are subsequently studied horizontally for their illocutionary meaning and perlocutionary reception.

The horizontal analysis follows a "pragmatic perspective, [whereby] language use and language users in interaction are primary, as opposed to language as a system of signs or a set of rules. The pragmatic perspective scrutinises neither just individual words nor sentences nor even isolated texts, but rather hold speech events or language games in real social contexts, considering both the present state of affairs and its connectedness with prior and succeeding action", according to Bublitz and Norrick (2011, p4).

Following the terminology introduced by Rühlemann and Aijmer (2015), the authors employ the term "*vertical*" to refer to occurrences of node words identified across the studied dialogues in a vertical direction and the term "*horizontal*" when several lines of text are analysed for their pragmatic function.

Fig. 1 depicts the different approach in vertical and horizontal corpus analyses.

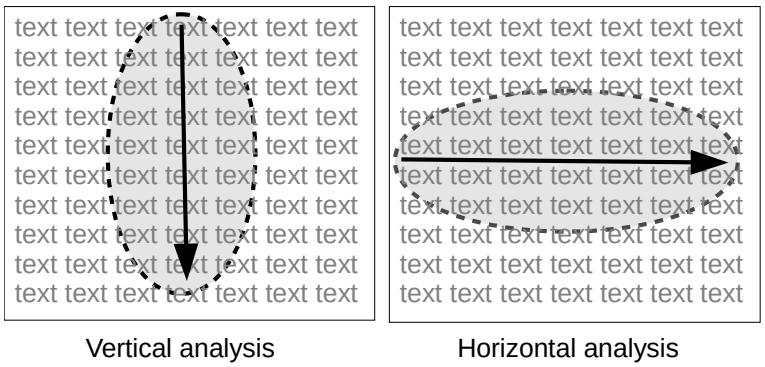

figure 1: Vertical and horizontal text analysis (Rühlemann and Aijmer, 2015)

Rühleman and Aijmer (2015, p55) refer to a number of researchers who have carried out "[c]orpus-based speech act studies [...] in which words and phrases, predetermined by the researcher to have pragmatic meaning, are searched in a corpus to identify speech acts (e.g., Adolphs, 2008; Aijmer, 1996, 2008; Jiang, 2006; Butler, 2008; Cheng, 2010)".

This research combines the search for predetermined words with a search for specific syntactic structures to classify speech acts as directives and commissives according to Searle's taxonomy of illocutionary acts (1975).

3 Identification of locutionary speech acts

In reference to Austin's original classification of speech acts (1962), Searle (1976, p10) proposed an alternative taxonomy with "basic categories of illocutionary acts". He defines the constructs of directives and commissives as follows:

- directives: "they are attempts […] by the speaker to get the hearer to do something" (Searle 1976, p11),
- commissives: they "commit the speaker (again in varying degrees) to some future course of action" (Searle 1976, p11).

After publishing his theoretical framework and stating examples of illocutionary acts, Searle provided his own semantic analysis for identifying speech acts (Searle and Vanderveken, 1985, pp179-217) on the basis of predefined lexical items, and although his taxonomy has since been used extensively by linguists (cf. Garcia McAllister, 2015, p34) no conclusive methodology has been developed for identifying speech acts classifiers. In the year 1996, Cohen (p385) saw a paradigm shift in that "although the process of defining and identifying speech acts has been going on since the 1960s, the last 15 years have marked a shift from intuitively based anecdotal approach to speech acts description to an empirical one [,...] encompassing both quantitative and qualitative approaches".

This research follows the more recent approach outlined by Qadir and Riloff (2011) who use "Lexical and Syntactic (LexSyn) Features, Speech Act Clue Features, and Semantic Features" to extract speech acts from a text corpus containing message board posts.

Upon applying the methodology proposed by Qadir and Riloff to the spoken corpus of bridge team communication and combining a number of "classifiers that can identify whether a sentence contains a speech act" (ibid, p748), a total of 3,172 speech acts clues for directives and commissives were identified in 2,218 utterances, leading to an overall speech act clue frequency of 29.7 clues per 1,000 tokens for the entire text corpus. The observed redundancy in the occurrence of speech acts clues (i.e. more than one clue per utterance) is based on the fact that speech acts clues are identified in the mixed-method manner described above.

The clues were subsequently annotated in the spoken corpus with pragmatic markers in line with research carried out by Garcia (2007), Rühlemann & O'Donnell (2012) and Stiles (1992) to allow for a vertical, quantitative analysis of the identified locutionary speech acts.

4 Results of the vertical analysis of locutionary speech acts

The annotated corpus was analysed for native and non-native speakers separately leading to significant differences between the number of speech act clues observed in L1 and L2 utterances. A Pearson's chi-squared test of association ($\chi 2$) carried out on these raw figures leads to $p<0.00$ so that the null hypothesis of no difference between the two frequencies must be rejected. The magnitude of the measured effect was

computed using the phi coefficient (Φ) which has been recommended by Grissom & Kim (2005) for naturalistic research on dichotomous variables (p249). On a scale from -1 to +1, phi results in 0.05 for directives and 0.04 for commissives where a value of 0 would denote no difference between the native and non-native speakers. Table 1 displays the individual clue values for both sub-corpora.

	Native speakers (L1)	**Non-native speakers (L2)**	**Effect size**
Speech act clue frequency per 1,000 word tokens	23.23	39.23	
Directives frequency	7.49 (479 clues)	13.48 (580 clues)	Φ=0.05*
Commissives frequency	15.73 (1,005 clues)	25.75 (1,108 clues)	Φ=0.04*
*Findings significant in $\chi 2$ test, α=0.01			

table 1: occurrences of speech act clues for directives and commissives

This vertical, corpus-based analysis on speech act clue frequencies has proved to be suitable for identifying a series of commissives and directives to be analysed horizontally. A between-group comparison of native and non-native speakers delivers divergent clue frequencies. Although the hypothesis testing leads to significant differences between native and non-native speakers, the phi coefficient points to rather marginal differences in the distribution of directives and commissives across both sub-corpora. It can thus be argued that at surface level, native and non-native speakers produce speech acts, including commissives and directives, at a similar rate.[4]

5 Categorising locutionary and illocutionary speech acts

After isolating a series of locutionary speech acts from the text corpus by means of a vertical text analysis, these speech acts are subsequently investigated for discrepancies between their locutionary and illocutionary meaning. It is understood that no difference between their locutionary and illocutionary meaning avoids ambiguities in the delivered

4 In their research on Speech Acts in Message Board Posts, Qadir and Riloff (2011) identify 159 directives and 261 commissives for each 1000 word tokens. Although the communicative setting is totally different, their findings coincide in a clear prevalence of directives over commissives.

message whereas a possible discrepancy puts more strain on the receiver to decode the message correctly.

For a qualitative analysis of possible differences between the literal meaning of an utterance and its illocutionary meaning, a dichotomous system has been chosen. This system grants a value of 0 to unambiguous utterances and a value of 1 to those utterances which can possibly lead to miscommunication due to a discrepancy between the locutionary act and its illocutionary meaning. Table 2 lists the number of possibly ambiguous directives and commissives uttered by native (L1) and non-native speakers (L2).

	L1	**L2**	**Effect size**
Possibly ambiguous utterances including directives (locutionary ≠ illocutionary act)	76 (15.8%**)	157 (27.0%**)	Φ=0.15*
Possibly ambiguous utterances including commissives (locutionary ≠ illocutionary act)	153 (15.2%**)	335 (30.0%**)	Φ=0.20*

*Findings significant in $\chi 2$ tests, $\alpha=0.01$
**Percentage of the clues listed in table 1

table 2: possibly ambiguous directives and commissives uttered by L1 and L2 speakers

The quantitative analysis of potentially ambiguous speech acts summarised in Table 2 reveals that language production by non-native speakers bears a much higher risk than messages uttered by native speakers. A calculation of the ratio of potentially ambiguous speech act production to the total number of clues leads to L2 figures which are approximately twice as high as for L1 speakers. With reference to the Osgood & Schramm communication model this means that in L2 communication, the effort to be made by the messages' receivers to decode them correctly can be assumed to cause a substantially higher risk of communicative disruptions.

Table 3 provides some examples for unambiguous and potentially ambiguous utterances. The lexical elements which can possibly lead to a misunderstanding are marked in *italics*.

The examples listed in Table 3 include lexical and syntactic features, speech act clue features, and semantic features, according to Qadir and Riloff (2011) classification categories as stated above. Lexical and syntactic feature include the use of personal

pronouns (e.g. 112_{L2}, 1246_{L2}, 10258_{L2}), future tense (e.g. 112_{L2}, 1246_{L2}, 10720_{L1}), modals (10258_{L1}, 64_{L2}, 1348_{L2}), infinitive verb phrases (e.g. 12250_{L1}, 1346_{L2}, 10161_{L1}), plan phrases (e.g. 1355_{L2}) and sentences beginning with modals or verbs, or with question words (e.g. 64_{L2}, 10258_{L1}, 10030_{L1}). Speech act clue words make use of the terms identified by Searle (1976) and Wierzbicka (1987), e.g. I suppose (10030_{L1}), and semantic features, e.g. to identify standardised nautical commands (as in utterance 15_{L2}).

unambiguous utterances (locutionary act = illocutionary meaning)	**utterances with a risk of ambiguity (locutionary act ≠ illocutionary meaning)**
(15_{L2}) Slow ahead, forty percent[1].**	(64_{L2}) *Can* you take over for a moment here?**
(112_{L2}) I will call him.*	(108_{L2}) *Can* you please call him to make *it* a *little bit* more?**
(1426_{L2}) So, I will tell you now our position.*	(1348_{L2}) But you know, you, you *should* use two radars.**
(5475_{L2}) Here then, please do the new waypoints, er, again.**	(1346_{L2}) I *only have to* switch the radar here on port side but so far it's ready, yeah.*
(10120_{L1}) Let's have a countdown now every half cable as opposed to every cable.***	(1355_{L2}) You're *gonna make sure* that you can start.**
(10258_{L1}) Shall I take a bearing now that she is stable?*	(10030_{L1}) When you are working this *I suppose you have to* keep because we are heading further south.**
(10720_{L1}) I will pull away.*	(10161_{L1}) I *have to* grab that now.*
(12250_{L1}) I tell you to get the bosun and tell him to get the anchors cleared away.**	(10424_{L1}) Let's stay *about there*, over.***

[1]standardised nautical command

*commissive, **directive, ***both commissive and directive

table 3: dichotomous system for differentiating utterances

The examples listed in the right column also include the typical use of *hedging* (Hyland, 1998; Lakoff, 1972; Markkanen and Schröder, 1997) as a means of politeness to mitigate criticism (e.g. in utterances 1348_{L2}, 1355_{L2}, 10030_{L1}). Table 4 lists the most frequent hedges according to Diani (2015, p180). It can be seen that modal verbs (*may*, *might*, *would*) occupy a prominent position in bridge team communication whilst premodifying adverbs (*perhaps*, *somewhat*, *rather*) only play a minor role with the exception of *probably* which is predominantly used for hedging purposes by the observed L1 speakers.

	L1	**L2**	***Effect size***
seem	15	3	Φ=0.06*
perhaps	0	1	**
may	4	31	Φ=0.07*
might	29	4	Φ=0.08*
probably	52	5	Φ=0.12*
would	95	68	Φ=0.05*
somewhat	0	1	**
rather	0	1	**

*Findings significant in $\chi 2$ tests, α=0.01; **findings not significant in $\chi 2$ tests, α=0.01

table 4: most frequent hedges in English as used by bridge teams

The total number of identified hedges is significantly higher in native speakers (n_{L1}=195; $frequency_{L1}$=0.13) than in non-native speakers (n_{L2}=114; $frequency_{L2}$=0.06). The latter tend to use a more direct wording in directives, as shown by the examples listed in Table 5. Here, non-native speakers use the simple present (2961_{L2}) and future tense (140_{L2}, 5835_{L2}) to direct team members in a matter-of-fact fashion whereas native speakers tend to use a more cautious wording which apparently leaves the decision with the communication partner: *you might have to* (7236_{L1}), *would you very quickly* (14261_{L1}).

Directives by native speakers (L1)	Directives by non-native speakers (L2)
(8587_{L1}) Now could you call me on the working channel, please, working channel please?	(108_{L2}) Can you please call him to make it a little bit more.
(7236_{L1}) You clear off these lads here, […] you might have to pass astern this guy.	(140_{L2}) So now you will ship away from the, here, at first, we will use the bow thruster, move to port.
(12169_{L1}) So you need to potentially remove that offset when that error or perceived error is removed.	(2961_{L2}) Okay, we move a little bit forward, and then you should start up.
(14261_{L1}) Would you very quickly get me the position on the radar, give me the range and bearing?	(5835_{L2}) After this you will send the sailing plan and the position report where we are now and where we are bound to.

table 5: directives uttered by native and non-native speakers

In directives, the use of hedging may cause misunderstanding as it apparently leaves the listener (i.e. receiver) with a choice which might not have been intended at all by the speaker (i.e. sender). Examples for this ambiguity can be seen in table 5 in utterances 2961_{L2}, 7236_{L1} and 12169_{L1}.

6 Results of the horizontal analysis of perlocutionary acts

So far, a vertical analysis has been carried out to isolate speech acts which potentially cause misunderstanding due to a discrepancy between their locutionary and illocutionary points. In the following, horizontal analyses are undertaken to determine if the communicative purpose of the isolated speech acts can be deemed to have been successful because the corresponding perlocutionary effect gives evidence of the fact that a message has been received and decoded correctly. For this purpose, all perlocutionary acts are categorised as follows: If the receiver closes the communicative loop by repeating substantial parts of the original message,[5] the speech act is considered successful and without any risk of miscommunication.

5 The loop is typically considered as closed when contents words and numbers are repeated by the receiver.

Assigned risk value	description	Clues for risk assessment
0	no risk of miscommunication	closing the communicative loop by partially repeating the original message *(27$_{L2}$) I stop the thrusters?* *(28$_{L2}$) Yeah, thruster stop please.* – *(4850$_{L2}$) Okay, so we will alter course to this, position approximately and also informate (sic.) the other two vessels about this, er, Ems and Jade.* *(4851$_{L2}$) That they can alter their course to this position.* – *(4947$_{L2}$) I suggest to decrease our speed.* *(4948$_{L2}$) Decrease, I will go to half ahead.* – *(11805$_{L1}$) Yeah, yeah, we will go for three hundred.* *(11806$_{L1}$) Three hundred?* *(11807$_{L1}$) Yeah.*
1	residual risk of miscommunication	answers limited to confirmation: yes, OK, good, I know, etc. *(15$_{L2}$) Slow ahead, forty percent.* *(16$_{L2}$) Yes.* – *(10718$_{L1}$) Yeah, I am not in the position back yet.* *(10719$_{L1}$) Okay.* *(10720$_{L1}$) I will pull away.* *(10721$_{L1}$) Do it.* – *(11964$_{L1}$) Let's come to 310.* *(11965$_{L1}$) Okay.* – *(15019$_{L1}$) We want to come around to port.* *(15020$_{L1}$) Good.*

Assigned risk value	description	Clues for risk assessment
2	high risk of miscommunication	no response, incorrect response or incoherent response *(6290$_{L2}$) I think you can use autopilot. [no response]* *(6291$_{L2}$) How often do you like the position?* *(6292$_{L2}$) Yeah, we can do, every six minutes.* – *(4905$_{L2}$) And where we have to go here?* *(4906$_{L2}$) Just a minute now, this chart. [incoherent response]* – *(4983$_{L2}$) Now, let's go.* *(4984$_{L2}$) So, I think, this is. [incoherent response]* – *(15009$_{L1}$) You can increase the engine slightly there too.* *(15010$_{L1}$) That is their plan, is it? [no response, refers to own observation instead]*

table 6: Examples for risk categories identified in perlocutionary acts

If, on the other hand, the receiver confirms the message without repeating parts of the original message, a residual risk of miscommunication exists as it is not clear whether the receiver has actually understood the message (i.e. decoded it correctly). The third category is formed by answers which clearly indicate that a communicative disruption has taken place. Table 6 lists the three categories including the clues used for their identification. It also states some examples for a better illustration of the used categories.

The assessment of possible risks in perlocutionary acts by means of the three discretionary categories of “no risk”, “residual risk” and “high risk” provides six pairs of observation for responses to directives and commissives. Given that the total word production (in tokens) varies between the two groups over an identical total recording time of 600 min each, the ratios of the identified risks to the occurrences of speech act clues are also given. Table 7 lists both the raw counts and the ratios (in percent). It also

includes the results of the hypothesis testing on differences between native and non-native speakers by means of Pearson's chi-squared tests of association.[6]

	Number of clues		X^2
	(L1)	(L2)	
Directives: no risk	37 (7.72*)	27 (4.66*)	p=0.05
Directives: residual risk	31 (6.47*)	44 (7.59*)	p=0.51
Directives: high risk	43 (8.98*)	63 (10.86*)	p=0.36
Commissives: no risk	85 (8.46*)	71 (6.41*)	p=0.09
Commissives: residual risk	114 (11.34*)	127 (11.46*)	p=0.94
Commissives: high risk	98 (9.75*)	130 (11.73*)	p=0.19

No findings significant in $\chi 2$ test, $\alpha=0.01$
*Percentage of the clues listed in table 1.

table 7: Risk assessment in perlocutionary acts

The data in table 7 leads to the conclusion that no significant differences must be assumed between the native and non-native speaker groups across any of the six risk categories. In other words, the likelihood of observing the categories of “no risk”, “residual risk” or “high risk” in the perlocutionary acts following directives and commissives is independent from the speakers’ mother tongue. The acceptance of the null hypothesis is corroborated by robust statistical results with a median X^2 test probability of p=0.27 on a sufficiently big sample size of n=3,172 speech acts clues.

6 In table 7 the total number of directives and commissives differs from the values in table 1 because some directives and commissives uttered by the sender did not receive any response from the receiver (e.g. in utterances 6290_{L2} and 15010_{L1}) while others are replied to more than once.

Even if both speaker groups present a similar distribution in their responses the risk of miscommunication remains quite high. In speech acts including directives, only 64 responses (L1 + L2) are given as an unambiguous closed-loop feedback whilst a total of 181 responses have a residual or high risk of misunderstanding. This means, that regardless of their mother tongue participants close the communicative loop in 35 percent of all speech acts only. According to Searle, by using directives the speaker "attempts [...] to get the hearer to do something" (1976, p11). A risk-minimising confirmation of a directive by the hearer in roughly one third of all responses is definitely not enough in the safety-critical environment on board a sea-going ship.

A similar situation can be ascertained for speech acts including commissives where a total of 156 closed-loop responses are opposed to 469 responses which include an inherent residual or high risk. As in the case of directives, just one third of all perlocutionary acts can be confirmed through a repetition of the original locutionary act. Here, it might be argued that commissives only "commit the speaker […] to some future course of action" (ibid, p11) without obliging the hearer to this particular future action. However, one important aim in using commissives at all is to inform other bridge team members about one's one thoughts to develop a shared mental model of the navigational situation. This is the very reason why Bridge Team Management training encourages participants to "think aloud" (Benedict, Kirchhoff, Gluch, Fischer, Schaub, & Baldauf 2015; Hederström, Kersandt, & Müller 2012). Again, observing closed-loop feedback in one third of all speech acts including commissives has to be considered a rather worrying ratio for a safety-critical work environment.

7. Findings and discussion

In her research on "performance of speech acts in workplace conversations and the teaching of communicative functions", Koester (2002, p167) refers to the complex nature of studying naturalistic speech acts: "Performing speech acts is a fairly complex phenomenon, which involves, according to Cohen (1996), sociocultural knowledge about when to perform a speech act and which one is appropriate in a given circumstance, as well as sociolinguistic knowledge regarding the actual linguistic realization of each speech act appropriate to the particular situation" (ibid, p168).

In bridge team communication the performance of speech acts fulfils the main purpose

of providing an effective, task-related communication with the overall objective to provide team members with sufficient information to steer a ship safely. Mariners shall use a coded language variety of English which aims to eliminate ambiguities by reducing lexical items and simplifying natural language structures. Theoretically, the "linguistic realization" (Koester, 2002, p168) of sociolinguistically appropriate speech acts in this domain should thus lead to a complete convergence of the illocutionary meaning with its corresponding locutionary act. On the other hand, adequate perlocutionary effects should be verified by using a closed-loop communication strategy. With reference to the Osgood & Schramm model (1954), risk-minimising maritime communication is achieved when messages are encoded correctly by senders, i.e. without discrepancies between their locutionary and illocutionary acts, and perlocutionary effects of decoded messages are verified by means of a closed-loop feedback strategy.

The intention of this research has been to determine to what extent native and non-native speakers of English actually employ the disambiguation approach stated above and in how far they are different from each other in this employment. Therefore, speech acts have firstly been identified in a vertical, corpus-based manner. This first step leads to markedly significant differences between directives and commissives produced by native and non-native speakers (see Table 1). After isolating directives and commissives vertically, a horizontal, pragmatic analysis of possible ambiguities has revealed that out of a total 2,218 utterances, 721 presented a difference between their locutionary and illocutionary points (see Table 2), thus causing an increased effort to decode the messages correctly with the inherent risk of misunderstanding.[7] Following the analysis of message senders, the responses by the message *receivers* has been scrutinised. Here, no significant differences were found between L1 and L2 participants. The analysis has revealed that only one third of all perlocutionary effects can be verified through closed-loop communication, while two thirds present a residual risk due to affirmative answers (e.g. *yes*, *good*, *okay*, etc.) or plainly incoherent responses (see Table 6).

According to Bublitz & Norrick "[p]ragmatics is fundamentally concerned with communicative action in any kind of context" (2011, p4). This *communicative action* is

7 The authors assume a model whereby indirect speech acts are typically harder to process than direct speech acts.

clearly visible in the maritime domain where work teams target specific professional issues to ensure the smooth operation of ships. The pragmatic analysis presented in this paper has studied naturalistic language *performance* in full-mission simulation exercises. Risks of miscommunication have been identified on both sides of the Osgood & Schramm communication model, and interestingly, significant differences between native and non-native speakers of English could only be ascertained in locutionary acts while responses to these resulted in non-significant results. The very similar risk patterns observed in L1 and L2 participants highlight the importance of a dedicated communication training once a sufficient language proficiency level has been achieved.

The Corpus Pragmatics approach has proved to be a valid tool for quantifying possible risks of miscommunication. The methodology and the research findings can be used to make bridge team communication more efficient and shipping safer by identifying inherent risks of miscommunication and raising future nautical officers' awareness in education and training sessions.

8. Limitations of the adopted methodology

The research was carried out on a special spoken corpus developed by the authors on the basis of recorded simulation exercises. Although exercises in full-mission simulation are highly realistic, an assumption of equality (Habermas, 1979) might be felt by participants which may not reflect the hierarchical structures on board real ships. Other variables leading to differences between simulations and real-life situations may include the multicultural nature of ship crews or their different language competence levels.

On the other hand, discretionary categorical variables have been assigned to the observed speech acts. By their very nature, the use of categorical variables for discrete values implies a simplification of the complex reality of speech acts.

To the authors' knowledge no ideal statistical test exists yet for measuring a dependent variable that consists of multiple categorical outcomes from speakers who contribute multiple data points (as is the case in this corpus). In this respect, the chosen chi-square test must be deemed anti-conservative because it assumes that all observations in the dataset are independent, which may give the analysis more power than is technically warranted by the dependence between data points from the same speaker or the same dialogue.

Lastly, the research has only considered verbal communication, thus disregarding any non-verbal communication like gestures which might have contributed to a disambiguation of possibly ambiguous speech acts.

9. Future research

Future research might look at the identified risks in more detail. Typical linguistic patterns might arise in highly risky speech acts which can be used for an improved communication training of future nautical officers. The applied methodology can also be used to study real-life situations on board sea-going ship and examine the differences among speakers in accordance with their general language competence levels and nationalities.

References

Adolphs, S. (2008). *Corpus and Context: Investigating Pragmatic Functions in Spoken Discourse*. Amsterdam and Philadelphia: John Benjamins.

Aijmer, K. (1996). *Conversational Routines in English: Convention and Creativity*. London: Longman.

Aijmer, K. (2008). At the interface between grammar and discourse: A corpus-based study of some pragmatic markers. In J. Romero-Trillo (Ed.), *Pragmatics and Corpus Linguistics: A Mutualistic Entente*, Berlin: Walter de Gruyter, 11–36.

Austin, J. L. (1962, 1975). *How to do things with words*. Oxford: Oxford University Press.

Bach, K. (1990). Communicative intentions, plan recognition, and pragmatics: Comments on

Bach, K., & Harnish, R. (1979). *Linguistic communication and speech acts*. London: MIT Press.

Baker, P., Ellece, S. (2011). *Key Terms in Discourse Analysis*. London: Bloomsbury.

Balmat, J. F., Lafont, F., Maifret, R., & Pessel, N. (2011). A decision-making system to maritime risk assessment. *Ocean Engineering*, *38(1)*, 171-176.

Benedict, K., Kirchhoff, M., Gluch, M., Fischer, S., Schaub, M., & Baldauf, M. (2015). Simulation-Augmented Methods for Manoeuvring Support-On-Board Ships and from the Shore. In Weintrit. A., Neumann. T. (Eds). *Information, Communication and Environment: Marine Navigation and Safety of Sea Transportation*, Taylor & Francis: London, 107-116.

Bocanegra-Valle, A. (2011). The Language of Seafaring: Standardized Conventions and Discursive Features in Speech Communications. *European Journal of English Studies, 11(1)*, 35-53.

Bocanegra-Valle, A. (2013). Maritime English. In C.A. Chapelle. *The Encyclopedia of Applied Linguistics*. New Jersey, Wiley. 3579-3583.

Bublitz, W., Norrick, N.R. (2011). Introduction: The burgeoning field of pragmatics. In W. Bublitz and N. R. Norrick (Eds.), *Foundations of Pragmatics, vol. 1 of Handbooks of Pragmatics*, Berlin: Mouton de Gruyter, 1–20.

Butler, C. (2008). The subjectivity of basically in British English: A corpus-based study. In J. Romero-Trillo (Ed.), *Pragmatics and Corpus Linguistics: A Mutualistic Entente*, Berlin: Walter de Gruyter, 37–63.

Chauvin, C., Lardjane, S., Morel, G., Clostermann, J. P., & Langard, B. (2013). Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. *Accident Analysis & Prevention*, *59*, 26-37.

Chawla, P. (2015). Happy talk: verbal communications and effective navigation. In The Nautical Institute (Eds.). *The Navigator, 08*, 4-5. Internet resource available at http://www.nautinst.org/en/Publications/the-navigator, viewed on 8 October 2017.

Cheng, S.W. (2010). A corpus-based approach to the study of speech act of thanking. *Concentric: Studies in Linguistics 36 (2),* 257–274.

Cohen, A.D. (1996). Speech acts. S.L. In S. McKay & N.H. Hornberger (Eds). *Sociolinguistics and Language Teaching*. Cambridge: Cambridge University Press, 383-420.

Davies, A. (2005). *A Glossary of Applied Linguistics*. Edinburgh: Edinburgh University Press.

de la Campa Portela, R. (2005). Maritime casualties analysis as a tool to improve research about human factors on maritime environment. *Journal of Maritime Research*, *2(2)*, 3-18.

Diani, G. (2015). Politeness. In K. Aijmer & C. Rühlmann (Eds). *Corpus Pragmatics. A handbook*. Cambridge: Cambridge University Press, 169-191.

Fasold, R. W. (1990). *Introduction to Sociolinguistics, vol. 2: The Sociolinguistics of Language*. Oxford: Blackwell.

Garcia, P. (2007). Pragmatics in academic contexts: A spoken corpus study. In M. C. Campoy and M. J. Luzón (Eds.), *Spoken Corpora in Applied Linguistics*. Bern: Peter Lang. 97-128.

Garcia McAllister, P. (2015). Speech acts: a synchronic perspective. In K. Aijmer & C. Rühlmann (Eds.). *Corpus Pragmatics. A handbook*. Cambridge: Cambridge University Press, 29-51.

Grissom, R.J., Kim, J.J. (2005). *Effect sizes for research. Univariate and Multivariate Applications*. Hove: Taylor & Francis.

Gustafsson, M. (2004). Simplification of Special Languages: a case study on Standard Marine Communication Phrases. In Hiltunen, R. (Ed.). *Approaches to style and discourse in English*. Osaka: Osaka University Press.

Habermas, J. (1979). *Communication and the evolution of society (Vol. 29)*, Boston: Beacon Press.

Hederström, H., Kersandt, D., & Müller, B. (2012). Task-oriented structure of the navigation process and quality control of its properties by a nautical task management monitor (NTMM). *European Journal of Navigation*, *10*(3), 1-20.

Holmes, J. (2005). Story-telling at work: a complex discursive resource for integrating personal, professional and social identities. *Discourse Studies*, *7(6)*, 671-700.

Hoover, J. (2002). *Effective Small Group and Team Communication*. Belmont: Wadsworth.

Horck, J. (2004). An analysis of decision-making processes in multicultural maritime scenarios. *Maritime policy & management*, *31(1)*, 15-29.

Hyland, K. (1998). *Hedging in Scientific Research Articles*. Amsterdam and Philadelphia: John Benjamins.

International Maritime Organization. (1978). *Standard Marine Navigational Vocabulary (SMNV)*. London: International Maritime Organization.

International Maritime Organization (2001). *Standard Marine Communication Phrases (SMCP)*. London: International Maritime Organization.

Jiang, X. (2006). Suggestions: What should ESL students know? *System, 34(1)*, 36–54.

John, P., Brooks, B. & Schriever, U. (2016). Linguistic measurement of cognitive load in maritime team communication by native and non-native speakers of English. Unpublished manuscript, Department of Maritime and Logistics Studies at Jade University of Applied Sciences, Elsfleth, Germany.

John, P., Brooks, B. & Schriever, U. (2017). Profiling maritime communication by non-native speakers: A quantitative comparison between the baseline and standard marine communication phraseology. *English for Specific Purposes, 47*, 1-14.

John, P., Brooks, B., Wand, C., & Schriever, U. (2013). Information density in bridge team communication and miscommunication – a quantitative approach to evaluate maritime communication. *WMU Journal of Maritime Affairs*, *12*, 229-244.

Jurkovič, V. (2015). Shared and Specific Features of Maritime English within the LSP Context. In Busch-Lauer, I. A. (Ed.). *Facetten der Fachsprachenvermittlung Englisch–Hands on ESP Teaching*, *5*. Berlin: Frank & Timme. 185-205.

Lent, B. (2013). 02:00 Communication Management: COM. In *Cybernetic Approach to Project Management*. Berlin: Springer. 337-364.

Koester, A.J. (2002). The performance of speech acts in workplace conversations and the teaching of communicative functions. *System 30(2)*, 167–184.

Kraut, R.E., Fish, R.S., Root, R.W., & Chalfonte, B.L. (1990). Informal communication

in organizations: Form, function, and technology. *Human reactions to technology: Claremont symposium on applied social psychology,* 145-199.

Lakoff, G. (1972). Hedges: A study of meaning criteria and the logic of fuzzy concepts. In P. M. Peranteau, J. N. Levi and G. C. Phares (Eds.), *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society.* Chicago: Chicago Linguistic Society. 183–228.

Leech, G. (1983). *Principles of Pragmatics*. London: Longman.

Levinson, S.C. (1983). *Pragmatics*. London: Longman.

Li, D. (2000). The pragmatics of making requests in the L2 workplace: A case study of language socialization. *Canadian Modern Language Review*, 57*(1)*, 58-87.

Lynch, O. H. (2002). Humorous communication: Finding a place for humour in communication research. *Communication theory*, *12(4)*, 423-445.

Markkanen, R., Schröder, H. (Eds). (1997). *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*. Berlin and New York: Mouton de Gruyter.

McCallum, M., Raby, M., Forsythe, A.M., Rothblum, A.M., Smith, M.W. (2000) Communications problems in marine casualties: Development and evaluation of investigation, reporting and analysis procedures. *Proceedings of the Human Factors and Ergonomics Society, Annual Meeting, 4,* 384-387

McCrae, C. (2009). Human factors at sea: common patterns of error in groundings and collisions. *Maritime Policy and Management*, *36(1)*, 21–38.

McQuail, D., & Windahl, S. (1993). *Communication models for the study of mass communications*. New York: Routledge.

Mey, J. (2001). *Pragmatics: An Introduction, second edition*. Oxford: Blackwell.

Noble, A. (2015). The IMO SMCP 15 years on: current perceptions and realistic recommendations, *Proceedings of the International Maritime English Conference*, Johor: International Maritime Lecturers' Association, 127-145. Internet resource available at:

http://www.pfri.uniri.hr/~bopri/IMEC_Proceedings/PDF/IMEC27.pdf, viewed on 8 October 2017.

Paltridge, B. (2006). *Discourse Analysis. An introduction*. London: Continuum.

Pritchard, B. & Kalogjera, D. (2000). On Some Features of Conversation in Maritime VHF Communication. In M. Coulthard, J. Cotterill & F. Rock (Eds.). *Dialogue Analysis VII: Working with Dialogue: Selected Papers from the 7th IADA Conference Birmingham 1999,* 185–194.

Pyne, R. & Koester, T. (2005) Methods and means for analysis of crew communication in the maritime domain. *The Archives of Transport, XVII, No. 3–4*, 193-208.

Qadir, A., & Riloff, E. (2011). Classifying sentences as speech acts in message board posts. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 748-758.

Rühlemann, C., Aijmer, K. (2015). *Corpus Pragmatics. A Handbook*. Cambridge: Cambridge University Press.

Rühlemann, C. and O'Donnell, M. B. (2012). Introducing a corpus of conversational narratives: Construction and annotation of the Narrative Corpus. *Corpus Linguistics and Linguistic Theory, 8(2)*, 313-350.

Schramm, W. (1954). How communication works. In W. Scramm (Ed.). *The Process and Effects of Mass Communication*. Urbana: University of Illinois Press. 3-26.

Searle, J.R. (1969). Speech acts: An essay in the philosophy of language, 626. Cambridge: Cambridge University Press.

Searle, J.R. (1975). A Classification of Illocutionary Acts. *Language in Society, 5(1)*, 1-23.

Searle, J.R. (1976). A taxonomy of illocutionary acts. In K. Günderson (Ed.). *Language, Mind, and Knowledge.* Minneapolis: University of Minnesota Press. 334-369.

Searle, J.R. & Vanderveken, D. (1985). Foundations of Illocutionary Logic. Cambridge: Cambridge University Press.

Stiles, W. B. (1992). Describing Talk: A Taxonomy of Verbal Response Modes. Newbury Park: Sage.

Velasquez, M., & Hester, P.T. (2013). An analysis of multi-criteria decision making methods. *International Journal of Operations Research*, *10*(2), 56-66.

11. Appendices

11.1 Appendix 1: Invitation to participate in linguistic research

International Maritime English in the doldrums:
A Comparison of Communication Strategies and Breakdowns of Maritime English by Native and Non-Native Bridge Team Members

The consent form will be signed by all trainees at the full-mission bridge simulator who wish to participate in the above mentioned linguistic research.

1. Invitation

You are invited to participate in a linguistic research project which analyses communication strategies used during the normal full-mission bridge simulator exercises. This study is being conducted in partial fulfilment of a PhD degree for Peter John under the supervision of Dr. Benjamin Brooks and Dr. Capt. Ulf Schriever. Peter John is a senior lecturer for Maritime English at the Maritime Faculty of Jade University of Applied Sciences (Germany). Dr. Benjamin Brooks is a Senior Research Fellow at the National Centre for Ports and Shipping, Dr. Cap. Ulf Schriever is a Lecturer of Maritime Training at the National Centre for Ports and Shipping.

2. What is the purpose of this study?

This study aims to introduce and validate a quantitative methodology to measure the information content in naturally occurring speech by the different members of a bridge team, including radio communication. Through the use of computational linguistics, the density of information is extracted for different conversation segments and related to the events taking place in the full-mission bridge simulations. By analysing audio-recorded and transcribed simulator exercises, a specific index suitable for evaluating individual and team communication performance will be developed.

3. Why have I been invited to participate?

You have been invited to participate by a random selection carried out by the researcher. Your involvement is totally voluntary, and there are no consequences whatsoever if you decide not to participate.

4. What will I be asked to do?

You will take part in the exercises at the full-mission bridge simulator which are part of your normal maritime training. Prior to this, you will be asked to fill in some details on your personal and professional background in an anonymous questionnaire and you will be audio-recorded. The recordings will later be transcribed and destroyed once the transcripts are completed. The transcripts and the information collected in the questionnaire will be included in a database and analysed linguistically.

5. Are there any possible benefits from participation in this study?

You may be provided with the research summary results. This will deliver a thorough analysis of bridge team communication beyond the scope of your own experience. The wider maritime community will benefit from an improved understanding of communication patterns during the normal operation of a ship and when resolving dangerous situations which may lead to marine accidents.

6. Are there any possible risks from participation in this study?

The participation does not involve any possible risk. You will carry out your normal training exercises at the full-mission bridge simulator without any modifications.

7. What if I change my mind during or after the study?

You are free to withdraw at any time, and you can do so without providing an explanation. Following your withdrawal, the questionnaire containing the collected information and the audio-recordings will be destroyed immediately without carrying out any transcription.

8. What will happen to the information when this study is over?

The original recordings (raw data) and all collected meta-data will be stored on a password-protected University of Tasmania computer until the transcription has been completed. Only the researcher and the supervisors will have access to the recordings. The transcripts and the collected meta-data will be included in an anonymous form in a database which will be stored on a password-protected University of Tasmania computer for 5 years from the date of first publication. Afterwards, the database will be archived on a password-protected computer at the researcher's premises, the transcripts and all

collected-meta data will be destroyed. All collected data will be treated in a confidential manner.

9. How will the results of the study be published?

The recordings will be transcribed and included in a database where the utterances are analysed using a Part-of-Speech (POS) tagging system to identify different word classes (e.g. nouns, verbs, adjectives, adverb). By applying descriptive and inferential statistics, typical speech patterns can be identified and linked to the meta-data collected. This allows for a quantitative analysis of differences between speakers of different mother tongues, different genders, etc. You will not be identifiable in the publication of the results.

10. What if I have questions about this study?

In case of any questions you may have about this study, you may contact the researcher by e-mail (Peter.John@utas.edu.au) or by phone (+49.44.....).

This study has been approved by the Tasmanian Social Sciences Human Research Ethics Committee. If you have concerns or complaints about the conduct of this study, please contact the Executive Officer of the HREC (Tasmania) Network on (03) 6226 7479 or email human.ethics@utas.edu.au . The Executive Officer is the person nominated to receive complaints from research participants. Please quote ethics reference number H0013035.

This information sheet is for you to keep. If you wish to participate in the research, you need to sign a written consent form.

11.2 Appendix 2: Participant consent form

International Maritime English in the doldrums:
A Comparison of Communication Strategies and Breakdowns of Maritime English by Native and Non-Native Bridge Team Members

The consent form has to be signed by all trainees at the full-mission bridge simulator who wish to participate in the above mentioned research.

1. I agree to take part in the research study named above.
2. I have read and understood the Information Sheet for this study.
3. The nature and possible effects of the study have been explained to me.
4. I understand that the study involves being audio-recorded during the training exercises carried out at the full-mission bridge simulator. The recordings will be transcribed in an anonymous form. The transcripts will be used for analysing communication patterns. The original recordings will be destroyed once the transcriptions have been finished.
5. I understand that participation does not involve any foreseeable risks.
6. I understand that all research data will be securely stored on the University of Tasmania premises for five years from the publication of the study results, and will then be destroyed unless I give permission for my data to be archived.

   I agree to have my study data archived.

   Yes □ No □
7. Any questions that I have asked have been answered to my satisfaction.
8. I understand that the researcher(s) will maintain confidentiality and that any information I supply to the researcher(s) will be used only for the purposes of the research.
9. I understand that the results of the study will be published so that I cannot be identified as a participant.

Yes □ No □

10. I understand that my participation is voluntary and that I may withdraw at any time without any effect.

I understand that I will not be able to withdraw my data after completing the exercise as the orginal recordings will be destroyed once they have been transcribed. The transcripts will be anonymous and cannot be linked to individual participants.

Participant's name: _____

Participant's signature: _____

Date: _____

Statement by Investigator

I have explained the project and the implications of participation in it to this volunteer and I believe that the consent is informed and that he/she understands the implications of participation.

If the Investigator has not had an opportunity to talk to participants prior to them participating, the following must be ticked.

The participant has received the Information Sheet where my details have been provided so participants have had the opportunity to contact me prior to consenting to participate in this project.

Investigator's name: Peter John

Investigator's signature: _____

Date: 28 January 2013

11.3 Appendix 3: Survey of participants' demographic data

International Maritime English in the doldrums:
A Comparison of Communication Strategies and Breakdowns of Maritime English by Native and Non-Native Bridge Team Members

Demographic data

Age

How old are you?

Answer:

Gender

Are you

☐ male or

☐ female?

Professional background

Sea experience

Give the total time you have worked on board sea-going ships.

(in months, for example, enter 5 for 5 months and 14 for one year and 2 months)

Answer:

Position on board

What is the highest position you have held on board a sea-going ship?

☐	☐	☐	☐	☐	☐	☐	☐
Cadet	Able seaman	Helmsman	Third officer	Second officer	First officer	Shipmaster	Other (e.g. Navy)

Last work

When did you last work on board a sea-going ship?

(Give the year, e.g. 2012)

Answer:

Working language

Language

When did you start learning English? Give the year.

(If English is your mother tongue, give the year you were born).

Answer:

Rating of English skills

Rating of English skills: how would you rate your English skills overall?

☐	☐	☐	☐	☐
Excellent	Very good	Good	Satisfactory	Poor

Nationality

What is your nationality?

Mother tongue

What is your mother tongue?

Multinational crew

Of the total time you worked on board seagoing ships, how often did you sail with multinational CREWS where English was the working language?

(If you have no seagoing experience, mark “no answer”.)

☐	☐	☐	☐	☐	☐
Always	Mostly	Half of the time	Hardly ever	Never	No answer

Communication amongst officers

Of the total time you worked on board seagoing ships, what language was mostly spoken amongst the OFFICERS?

(If you have no seagoing experience, mark “no answer”).

☐	☐	☐	☐	☐
English	My mother tongue (if not English)	Another language I speak	Another language I do not speak	No answer